SEQUENCE LISTING

(1) GENERAL INFORMATION:

5

(i) APPLICANT:

10

(ii) TITLE OF INVENTION:

(iii) NUMBER OF SEQUENCES: 19

15

(iv) CORRESPONDENCE ADDRESS:
    (A)   ADDRESSEE: Foley, Hoag & Eliot LLP
    (B)   STREET: One Post Office Square
    (C)   CITY: Boston

20
    (D)   STATE: MA
    (E)   COUNTRY: US
    (F)   ZIP: 02109

(v) COMPUTER READABLE FORM:

25
    (A)   MEDIUM TYPE: Floppy disk
    (B)   COMPUTER: IBM PC compatible
    (C)   OPERATING SYSTEM: PC-DOS/MS-DOS
    (D)   SOFTWARE: PatentIn Release #1.0, Version #1.30

30 (vi) CURRENT APPLICATION DATA:
    (A)   APPLICATION NUMBER:
    (B)   FILING DATE:
    (C)   CLASSIFICATION:

35 (viii) ATTORNEY/AGENT INFORMATION:
    (A)   NAME: Vincent, Matthew P.
    (B)   REGISTRATION NUMBER: 36,709
    (C)   REFERENCE/DOCKET NUMBER: SUV003.04

40 (ix) TELECOMMUNICATION INFORMATION:
    (A)   TELEPHONE: 617-832-1000
    (B)   TELEFAX: 617-832-7000

45 (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
    (A)   LENGTH: 736 base pairs
    (B)   TYPE: nucleic acid

50
    (C)   STRANDEDNESS: single
    (D)   TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
AACNNCNNTN NATGGCACCC CCNCCCAACC TTTNNNCCNN NTAANCAAAA NNCCCCNTTT   60

NATACCCCCT NTAANANTTT TCCACCNNNC NNAAANNCCN CTGNANACNA NGNAAANCCN  120

TTTTTNAACC CCCCCCACCC GGAATTCCNA NTNNCCNCCC CCAAATTACA ACTCCAGNCC  180
```

60

```
AAAATTNANA NAATTGGTCC TAACCTAACC NATNGTTGTT ACGGTTTCCC CCCCCAAATA    240

CATGCACTGG CCCGAACACT TGATCGTTGC CGTTCCAATA AGAATAAATC TGGTCATATT    300

AAACAAGCCN AAAGCTTTAC AAACTGTTGT ACAATTAATG GGCGAACACG AACTGTTCGA    360

ATTCTGGTCT GGACATTACA AAGTGCACCA CATCGGATGG AACCAGGAGA AGGCCACAAC    420

CGTACTGAAC GCCTGGCAGA AGAAGTTCGC ACAGGTTGGT GGTTGGCGCA AGGAGTAGAG    480

TGAATGGTGG TAATTTTTGG TTGTTCCAGG AGGTGGATCG TCTGACGAAG AGCAAGAAGT    540

CGTCGAATTA CATCTTCGTG ACGTTCTCCA CCGCCAATTT GAACAAGATG TTGAAGGAGG    600

CGTCGAANAC GGACGTGGTG AAGCTGGGGG TGGTGCTGGG GGTGGCGGCG GTGTACGGGT    660

GGGTGGCCCA GTCGGGGCTG GCTGCCTTGG GAGTGCTGGT CTTNGCGNGC TNCNATTCGC    720

CCTATAGTNA GNCGTA                                                   736
```

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 107 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Xaa Pro Pro Pro Asn Tyr Asn Ser Xaa Pro Lys Xaa Xaa Xaa Leu Val
1               5                   10              15

Leu Thr Pro Xaa Val Val Thr Val Ser Pro Pro Lys Tyr Met His Trp
            20              25              30

Pro Glu His Leu Ile Val Ala Val Pro Ile Arg Ile Asn Leu Val Ile
        35              40              45

Leu Asn Lys Pro Lys Ala Leu Gln Thr Val Val Gln Leu Met Gly Glu
    50              55              60

His Glu Leu Phe Glu Phe Trp Ser Gly His Tyr Lys Val His His Ile
65              70              75              80

Gly Trp Asn Gln Glu Lys Ala Thr Thr Val Leu Asn Ala Trp Gln Lys
            85              90              95

Lys Phe Ala Gln Val Gly Gly Trp Arg Lys Glu
            100             105
```

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5187 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
GGGTCTGTCA CCCGGAGCCG GAGTCCCCGG CGGCCAGCAG CGTCCTCGCG AGCCGAGCGC       60

CCAGGCGCGC CCGGAGCCCG CGGCGGCGGC GGCAACATGG CCTCGGCTGG TAACGCCGCC      120

GGGGCCCTGG GCAGGCAGGC CGGCGGCGGG AGGCGCAGAC GGACCGGGGG ACCGCACCGC      180

GCCGCGCCGG ACCGGGACTA TCTGCACCGG CCCAGCTACT GCGACGCCGC CTTCGCTCTG      240

GAGCAGATTT CCAAGGGGAA GGCTACTGGC CGGAAAGCGC CGCTGTGGCT GAGAGCGAAG      300

TTTCAGAGAC TCTTATTTAA ACTGGGTTGT TACATTCAAA AGAACTGCGG CAAGTTTTTG      360

GTTGTGGGTC TCCTCATATT TGGGGCCTTC GCTGTGGGAT TAAAGGCAGC TAATCTCGAG      420

ACCAACGTGG AGGAGCTGTG GGTGGAAGTT GGTGGACGAG TGAGTCGAGA ATTAAATTAT      480

ACCCGTCAGA AGATAGGAGA AGAGGCTATG TTTAATCCTC AACTCATGAT ACAGACTCCA      540

AAAGAAGAAG GCGCTAATGT TCTGACCACA GAGGCTCTCC TGCAACACCT GGACTCAGCA      600

CTCCAGGCCA GTCGTGTGCA CGTCTACATG TATAACAGGC AATGGAAGTT GGAACATTTG      660

TGCTACAAAT CAGGGGAACT TATCACGGAG ACAGGTTACA TGGATCAGAT AATAGAATAC      720

CTTTACCCTT GCTTAATCAT TACACCTTTG GACTGCTTCT GGGAAGGGGC AAAGCTACAG      780

TCCGGGACAG CATACCTCCT AGGTAAGCCT CCTTTACGGT GGACAAACTT TGACCCCTTG      840

GAATTCCTAG AAGAGTTAAA GAAAATAAAC TACCAAGTGG ACAGCTGGGA GGAAATGCTG      900

AATAAAGCCG AAGTTGGCCA TGGGTACATG GACCGGCCTT GCCTCAACCC AGCCGACCCA      960

GATTGCCCTG CCACAGCCCC TAACAAAAAT TCAACCAAAC CTCTTGATGT GGCCCTTGTT     1020

TTGAATGGTG GATGTCAAGG TTTATCCAGG AAGTATATGC ATTGGCAGGA GGAGTTGATT     1080

GTGGGTGGTA CCGTCAAGAA TGCCACTGGA AAACTTGTCA GCGCTCACGC CCTGCAAACC     1140

ATGTTCCAGT TAATGACTCC CAAGCAAATG TATGAACACT TCAGGGGCTA CGACTATGTC     1200

TCTCACATCA ACTGGAATGA AGACAGGGCA GCCGCCATCC TGGAGGCCTG GCAGAGGACT     1260
TACGTGGAGG TGGTTCATCA AAGTGTCGCC CCAAACTCCA CTCAAAAGGT GCTTCCCTTC     1320

ACAACCACGA CCCTGGACGA CATCCTAAAA TCCTTCTCTG ATGTCAGTGT CATCCGAGTG     1380

GCCAGCGGCT ACCTACTGAT GCTTGCCTAT GCCTGTTTAA CCATGCTGCG CTGGGACTGC     1440

TCCAAGTCCC AGGGTGCCGT GGGGCTGGCT GGCGTCCTGT TGGTTGCGCT GTCAGTGGCT     1500

GCAGGATTGG GCCTCTGCTC CTTGATTGGC ATTTCTTTTA ATGCTGCGAC AACTCAGGTT     1560

TTGCCGTTTC TTGCTCTTGG TGTTGGTGTG GATGATGTCT TCCTCCTGGC CCATGCATTC     1620

AGTGAAACAG GACAGAATAA GAGGATTCCA TTTGAGGACA GGACTGGGGA GTGCCTCAAG     1680
```

```
CGCACCGGAG CCAGCGTGGC CCTCACCTCC ATCAGCAATG TCACCGCCTT CTTCATGGCC    1740

GCATTGATCC CTATCCCTGC CCTGCGAGCG TTCTCCCTCC AGGCTGCTGT GGTGGTGGTA    1800

TTCAATTTTG CTATGGTTCT GCTCATTTTT CCTGCAATTC TCAGCATGGA TTTATACAGA    1860

CGTGAGGACA GAAGATTGGA TATTTTCTGC TGTTTCACAA GCCCCTGTGT CAGCAGGGTG    1920

ATTCAAGTTG AGCCACAGGC CTACACAGAG CCTCACAGTA ACACCCGGTA CAGCCCCCCA    1980

CCCCCATACA CCAGCCACAG CTTCGCCCAC GAAACCCATA TCACTATGCA GTCCACCGTT    2040

CAGCTCCGCA CAGAGTATGA CCCTCACACG CACGTGTACT ACACCACCGC CGAGCCACGC    2100

TCTGAGATCT CTGTACAGCC TGTTACCGTC ACCCAGGACA ACCTCAGCTG TCAGAGTCCC    2160

GAGAGCACCA GCTCTACCAG GGACCTGCTC TCCCAGTTCT CAGACTCCAG CCTCCACTGC    2220

CTCGAGCCCC CCTGCACCAA GTGGACACTC TCTTCGTTTG CAGAGAAGCA CTATGCTCCT    2280

TTCCTCCTGA AACCCAAAGC CAAGGTTGTG GTAATCCTTC TTTTCCTGGG CTTGCTGGGG    2340

GTCAGCCTTT ATGGGACCAC CCGAGTGAGA GACGGGCTGG ACCTCACGGA CATTGTTCCC    2400

CGGGAAACCA GAGAATATGA CTTCATAGCT GCCCAGTTCA AGTACTTCTC TTTCTACAAC    2460

ATGTATATAG TCACCCAGAA AGCAGACTAC CCGAATATCC AGCACCTACT TTACGACCTT    2520

CATAAGAGTT TCAGCAATGT GAAGTATGTC ATGCTGGAGG AGAACAAGCA ACTTCCCCAA    2580

ATGTGGCTGC ACTACTTTAG AGACTGGCTT CAAGGACTTC AGGATGCATT TGACAGTGAC    2640

TGGGAAACTG GGAGGATCAT GCCAAACAAT TATAAAAATG GATCAGATGA CGGGGTCCTC    2700

GCTTACAAAC TCCTGGTGCA GACTGGCAGC CGAGACAAGC CCATCGACAT TAGTCAGTTG    2760

ACTAAACAGC GTCTGGTAGA CGCAGATGGC ATCATTAATC CGAGCGCTTT CTACATCTAC    2820

CTGACCGCTT GGGTCAGCAA CGACCCTGTA GCTTACGCTG CCTCCCAGGC CAACATCCGG    2880

CCTCACCGGC CGGAGTGGGT CCATGACAAA GCCGACTACA TGCCAGAGAC CAGGCTGAGA    2940

ATCCCAGCAG CAGAGCCCAT CGAGTACGCT CAGTTCCCTT TCTACCTCAA CGGCCTACGA    3000

GACACCTCAG ACTTTGTGGA AGCCATAGAA AAAGTGAGAG TCATCTGTAA CAACTATACG    3060

AGCCTGGGAC TGTCCAGCTA CCCCAATGGC TACCCCTTCC TGTTCTGGGA GCAATACATC    3120

AGCCTGCGCC ACTGGCTGCT GCTATCCATC AGCGTGGTGC TGGCCTGCAC GTTTCTAGTG    3180

TGCGCAGTCT TCCTCCTGAA CCCCTGGACG GCCGGGATCA TTGTCATGGT CCTGGCTCTG    3240

ATGACCGTTG AGCTCTTTGG CATGATGGGC CTCATTGGGA TCAAGCTGAG TGCTGTGCCT    3300

GTGGTCATCC TGATTGCATC TGTTGGCATC GGAGTGGAGT TCACCGTCCA CGTGGCTTTG    3360

GCCTTTCTGA CAGCCATTGG GGACAAGAAC CACAGGGCTA TGCTCGCTCT GGAACACATG    3420

TTTGCTCCCG TTCTGGACGG TGCTGTGTCC ACTCTGCTGG GTGTACTGAT GCTTGCAGGG    3480

TCCGAATTTG ATTTCATTGT CAGATACTTC TTTGCCGTCC TGGCCATTCT CACCGTCTTG    3540

GGGGTTCTCA ATGGACTGGT TCTGCTGCCT GTCCTCTTAT CCTTCTTTGG ACCGTGTCCT    3600
```

```
GAGGTGTCTC CAGCCAATGG CCTAAACCGA CTGCCCACTC CTTCGCCTGA GCCGCCTCCA    3660

AGTGTCGTCC GGTTTGCCGT GCCTCCTGGT CACACGAACA ATGGGTCTGA TTCCTCCGAC    3720

TCGGAGTACA GCTCTCAGAC CACGGTGTCT GGCATCAGTG AGGAGCTCAG GCAATACGAA    3780

GCACAGCAGG GTGCCGGAGG CCCTGCCCAC CAAGTGATTG TGGAAGCCAC AGAAAACCCT    3840

GTCTTTGCCC GGTCCACTGT GGTCCATCCG GACTCCAGAC ATCAGCCTCC CTTGACCCCT    3900

CGGCAACAGC CCCACCTGGA CTCTGGCTCC TTGTCCCCTG GACGGCAAGG CCAGCAGCCT    3960

CGAAGGGATC CCCCTAGAGA AGGCTTGCGG CCACCCCCCT ACAGACCGCG CAGAGACGCT    4020

TTTGAAATTT CTACTGAAGG GCATTCTGGC CCTAGCAATA GGGACCGCTC AGGGCCCCGT    4080

GGGGCCCGTT CTCACAACCC TCGGAACCCA ACGTCCACCG CCATGGGCAG CTCTGTGCCC    4140

AGCTACTGCC AGCCCATCAC CACTGTGACG GCTTCTGCTT CGGTGACTGT TGCTGTGCAT    4200

CCCCCGCCTG GACCTGGGCG CAACCCCCGA GGGGGGCCCT GTCCAGGCTA TGAGAGCTAC    4260

CCTGAGACTG ATCACGGGGT ATTTGAGGAT CCTCATGTGC CTTTTCATGT CAGGTGTGAG    4320

AGGAGGGACT CAAAGGTGGA GGTCATAGAG CTACAGGACG TGGAATGTGA GGAGAGGCCG    4380

TGGGGGAGCA GCTCCAACTG AGGGTAATTA AAATCTGAAG CAAAGAGGCC AAAGATTGGA    4440

AAGCCCCGCC CCCACCTCTT TCCAGAACTG CTTGAAGAGA ACTGCTTGGA ATTATGGGAA    4500

GGCAGTTCAT TGTTACTGTA ACTGATTGTA TTATTKKGTG AAATATTTCT ATAAATATTT    4560

AARAGGTGTA CACATGTAAT ATACATGGAA ATGCTGTACA GTCTATTTCC TGGGGCCTCT    4620

CCACTCCTGC CCCAGAGTGG GGAGACCACA GGGGCCCTTT CCCCTGTGTA CATTGGTCTC    4680

TGTGCCACAA CCAAGCTTAA CTTAGTTTTA AAAAAAATCT CCCAGCATAT GTCGCTGCTG    4740

CTTAAATATT GTATAATTTA CTTGTATAAT TCTATGCAAA TATTGCTTAT GTAATAGGAT    4800

TATTTGTAAA GGTTTCTGTT TAAAATATTT TAAATTTGCA TATCACAACC CTGTGGTAGG    4860

ATGAATTGTT ACTGTTAACT TTTGAACACG CTATGCGTGG TAATTGTTTA ACGAGCAGAC    4920

ATGAAGAAAA CAGGTTAATC CCAGTGGCTT CTCTAGGGGT AGTTGTATAT GGTTCGCATG    4980

GGTGGATGTG TGTGTGCATG TGACTTTCCA ATGTACTGTA TTGTGGTTTG TTGTTGTTGT    5040

TGCTGTTGTT GTTCATTTTG GTGTTTTTGG TTGCTTTGTA TGATCTTAGC TCTGGCCTAG    5100

GTGGGCTGGG AAGGTCCAGG TCTTTTTCTG TCGTGATGCT GGTGGAAAGG TGACCCCAAT    5160

CATCTGTCCT ATTCTCTGGG ACTATTC                                       5187
```

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1311 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Val Ala Pro Asp Ser Glu Ala Pro Ser Asn Pro Arg Ile Thr Ala
1               5                   10                  15

Ala His Glu Ser Pro Cys Ala Thr Glu Ala Arg His Ser Ala Asp Leu
            20                  25                  30

Tyr Ile Arg Thr Ser Trp Val Asp Ala Ala Leu Ala Leu Ser Glu Leu
        35                  40                  45

Glu Lys Gly Asn Ile Glu Gly Gly Arg Thr Ser Leu Trp Ile Arg Ala
    50                  55                  60

Trp Leu Gln Glu Gln Leu Phe Ile Leu Gly Cys Phe Leu Gln Gly Asp
65                  70                  75                  80

Ala Gly Lys Val Leu Phe Val Ala Ile Leu Val Leu Ser Thr Phe Cys
            85                  90                  95

Val Gly Leu Lys Ser Ala Gln Ile His Thr Arg Val Asp Gln Leu Trp
            100                 105                 110

Val Gln Glu Gly Gly Arg Leu Glu Ala Glu Leu Lys Tyr Thr Ala Gln
        115                 120                 125

Ala Leu Gly Glu Ala Asp Ser Ser Thr His Gln Leu Val Ile Gln Thr
    130                 135                 140

Ala Lys Asp Pro Asp Val Ser Leu Leu His Pro Gly Ala Leu Leu Glu
145                 150                 155                 160

His Leu Lys Val Val His Ala Ala Thr Arg Val Thr Val His Met Tyr
            165                 170                 175

Asp Ile Glu Trp Arg Leu Lys Asp Leu Cys Tyr Ser Pro Ser Ile Pro
            180                 185                 190

Asp Phe Glu Gly Tyr His His Ile Glu Ser Ile Ile Asp Asn Val Ile
        195                 200                 205

Pro Cys Ala Ile Ile Thr Pro Leu Asp Cys Phe Trp Glu Gly Ser Lys
    210                 215                 220

Leu Leu Gly Pro Asp Tyr Pro Ile Tyr Val Pro His Leu Lys His Lys
225                 230                 235                 240

Leu Gln Trp Thr His Leu Asn Pro Leu Glu Val Val Glu Glu Val Lys
            245                 250                 255

Lys Leu Lys Phe Gln Phe Pro Leu Ser Thr Ile Glu Ala Tyr Met Lys
            260                 265                 270

Arg Ala Gly Ile Thr Ser Ala Tyr Met Lys Lys Pro Cys Leu Asp Pro
    275                 280                 285

Thr Asp Pro His Cys Pro Ala Thr Ala Pro Asn Lys Lys Ser Gly His
    290                 295                 300
```

```
Ile Pro Asp Val Ala Ala Glu Leu Ser His Gly Cys Tyr Gly Phe Ala
305              310              315              320

Ala Ala Tyr Met His Trp Pro Glu Gln Leu Ile Val Gly Gly Ala Thr
             325              330              335

Arg Asn Ser Thr Ser Ala Leu Arg Lys Ala Arg Xaa Leu Gln Thr Val
             340              345              350

Val Gln Leu Met Gly Glu Arg Glu Met Tyr Glu Tyr Trp Ala Asp His
             355              360              365

Tyr Lys Val His Gln Ile Gly Trp Asn Gln Glu Lys Ala Ala Ala Val
    370              375              380

Leu Asp Ala Trp Gln Arg Lys Phe Ala Ala Glu Val Arg Lys Ile Thr
385              390              395              400

Thr Ser Gly Ser Val Ser Ser Ala Tyr Ser Phe Tyr Pro Phe Ser Thr
             405              410              415

Ser Thr Leu Asn Asp Ile Leu Gly Lys Phe Ser Glu Val Ser Leu Lys
             420              425              430

Asn Ile Ile Leu Gly Tyr Met Phe Met Leu Ile Tyr Val Ala Val Thr
         435              440              445

Leu Ile Gln Trp Arg Asp Pro Ile Arg Ser Gln Ala Gly Val Gly Ile
    450              455              460

Ala Gly Val Leu Leu Leu Ser Ile Thr Val Ala Ala Gly Leu Gly Phe
465              470              475              480

Cys Ala Leu Leu Gly Ile Pro Phe Asn Ala Ser Ser Thr Gln Ile Val
             485              490              495

Pro Phe Leu Ala Leu Gly Leu Gly Val Gln Asp Met Phe Leu Leu Thr
             500              505              510

His Thr Tyr Val Glu Gln Ala Gly Asp Val Pro Arg Glu Glu Arg Thr
    515              520              525

Gly Leu Val Leu Lys Lys Ser Gly Leu Ser Val Leu Leu Ala Ser Leu
    530              535              540

Cys Asn Val Met Ala Phe Leu Ala Ala Ala Leu Leu Pro Ile Pro Ala
545              550              555              560

Phe Arg Val Phe Cys Leu Gln Ala Ala Ile Leu Leu Leu Phe Asn Leu
             565              570              575

Gly Ser Ile Leu Leu Val Phe Pro Ala Met Ile Ser Leu Asp Leu Arg
         580              585              590

Arg Arg Ser Ala Ala Arg Ala Asp Leu Leu Cys Cys Leu Met Pro Glu
         595              600              605

Ser Pro Leu Pro Lys Lys Lys Ile Pro Glu Arg Ala Lys Thr Arg Lys
    610              615              620

Asn Asp Lys Thr His Arg Ile Asp Thr Thr Arg Gln Pro Leu Asp Pro
    625              630              635              640
```

```
Asp Val Ser Glu Asn Val Thr Lys Thr Cys Cys Leu Ser Val Ser Leu
            645                 650             655

Thr Lys Trp Ala Lys Asn Gln Tyr Ala Pro Phe Ile Met Arg Pro Ala
            660                 665             670

Val Lys Val Thr Ser Met Leu Ala Leu Ile Ala Val Ile Leu Thr Ser
            675                 680             685

Val Trp Gly Ala Thr Lys Val Lys Asp Gly Leu Asp Leu Thr Asp Ile
            690                 695             700

Val Pro Glu Asn Thr Asp Glu His Glu Phe Leu Ser Arg Gln Glu Lys
705                 710                 715             720

Tyr Phe Gly Phe Tyr Asn Met Tyr Ala Val Thr Gln Gly Asn Phe Glu
            725                 730             735

Tyr Pro Thr Asn Gln Lys Leu Leu Tyr Glu Tyr His Asp Gln Phe Val
            740                 745             750

Arg Ile Pro Asn Ile Ile Lys Asn Asp Asn Gly Gly Leu Thr Lys Phe
            755                 760             765

Trp Leu Ser Leu Phe Arg Asp Trp Leu Leu Asp Leu Gln Val Ala Phe
            770                 775             780

Asp Lys Glu Val Ala Ser Gly Cys Ile Thr Gln Glu Tyr Trp Cys Lys
785                 790                 795             800

Asn Ala Ser Asp Glu Gly Ile Leu Ala Tyr Lys Leu Met Val Gln Thr
                805                 810             815

Gly His Val Asp Asn Pro Ile Asp Lys Ser Leu Ile Thr Ala Gly His
            820                 825             830

Arg Leu Val Asp Lys Asp Gly Ile Ile Asn Pro Lys Ala Phe Tyr Asn
            835                 840             845

Tyr Leu Ser Ala Trp Ala Thr Asn Asp Ala Leu Ala Tyr Gly Ala Ser
            850                 855             860

Gln Gly Asn Leu Lys Pro Gln Pro Gln Arg Trp Ile His Ser Pro Glu
865                 870                 875             880

Asp Val His Leu Glu Ile Lys Lys Ser Ser Pro Leu Ile Tyr Thr Gln
                885                 890             895

Leu Pro Phe Tyr Leu Ser Gly Leu Ser Asp Thr Xaa Ser Ile Lys Thr
            900                 905             910

Leu Ile Arg Ser Val Arg Asp Leu Cys Leu Lys Tyr Glu Ala Lys Gly
            915                 920             925

Leu Pro Asn Phe Pro Ser Gly Ile Pro Phe Leu Phe Trp Glu Gln Tyr
            930                 935             940

Leu Tyr Leu Arg Thr Ser Leu Leu Leu Ala Leu Ala Cys Ala Leu Ala
945                 950                 955             960

Ala Val Phe Ile Ala Val Met Val Leu Leu Leu Asn Ala Trp Ala Ala
                965                 970             975
```

Val Leu Val Thr Leu Ala Leu Ala Thr Leu Val Leu Gln Leu Leu Gly
      980   985   990

Val Met Ala Leu Leu Gly Val Lys Leu Ser Ala Met Pro Ala Val Leu
      995   1000   1005

Leu Val Leu Ala Ile Gly Arg Gly Val His Phe Thr Val His Leu Cys
     1010   1015   1020

Leu Gly Phe Val Thr Ser Ile Gly Cys Lys Arg Arg Arg Ala Ser Leu
1025   1030   1035   1040

Ala Leu Glu Ser Val Leu Ala Pro Val Val His Gly Ala Leu Ala Ala
         1045   1050   1055

Ala Leu Ala Ala Ser Met Leu Ala Ala Ser Glu Cys Gly Phe Val Ala
         1060   1065   1070

Arg Leu Phe Leu Arg Leu Leu Leu Asp Ile Val Phe Leu Gly Leu Ile
        1075   1080   1085

Asp Gly Leu Leu Phe Phe Pro Ile Val Leu Ser Ile Leu Gly Pro Ala
        1090   1095   1100

Ala Glu Val Arg Pro Ile Glu His Pro Glu Arg Leu Ser Thr Pro Ser
1105   1110   1115   1120

Pro Lys Cys Ser Pro Ile His Pro Arg Lys Ser Ser Ser Ser Ser Gly
        1125   1130   1135

Gly Gly Asp Lys Ser Ser Arg Thr Ser Lys Ser Ala Pro Arg Pro Cys
        1140   1145   1150

Ala Pro Ser Leu Thr Thr Ile Thr Glu Glu Pro Ser Ser Trp His Ser
        1155   1160   1165

Ser Ala His Ser Val Gln Ser Ser Met Gln Ser Ile Val Val Gln Pro
        1170   1175   1180

Glu Val Val Val Glu Thr Thr Thr Tyr Asn Gly Ser Asp Ser Ala Ser
1185   1190   1195   1200

Gly Arg Ser Thr Pro Thr Lys Ser Ser His Gly Gly Ala Ile Thr Thr
        1205   1210   1215

Thr Lys Val Thr Ala Thr Ala Asn Ile Lys Val Glu Val Val Thr Pro
        1220   1225   1230

Ser Asp Arg Lys Ser Arg Arg Ser Tyr His Tyr Tyr Asp Arg Arg Arg
        1235   1240   1245

Asp Arg Asp Glu Asp Arg Asp Arg Asp Arg Glu Arg Asp Arg Asp Arg
        1250   1255   1260

Asp Arg Asp Arg Asp Arg Asp Arg Asp Arg Asp Arg Asp Arg Asp Arg
1265   1270   1275   1280

Glu Arg Ser Arg Glu Arg Asp Arg Arg Asp Arg Tyr Arg Asp Glu Arg
        1285   1290   1295

Asp His Arg Ala Ser Pro Arg Glu Lys Arg Gln Arg Phe Trp Thr
        1300   1305   1310

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4434 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
CGAAACAAGA GAGCGAGTGA GAGTAGGGAG AGCGTCTGTG TTGTGTGTTG AGTGTCGCCC        60

ACGCACACAG GCGCAAAACA GTGCACACAG ACGCCCGCTG GGCAAGAGAG AGTGAGAGAG       120

AGAAACAGCG GCGCGCGCTC GCCTAATGAA GTTGTTGGCC TGGCTGGCGT GCCGCATCCA       180

CGAGATACAG ATACATCTCT CATGGACCGC GACAGCCTCC CACGCGTTCC GGACACACAC       240

GGCGATGTGG TCGATGAGAA ATTATTCTCG GATCTTTACA TACGCACCAG CTGGGTGGAC       300

GCCCAAGTGG CGCTCGATCA GATAGATAAG GGCAAAGCGC GTGGCAGCCG CACGGCGATC       360

TATCTGCGAT CAGTATTCCA GTCCCACCTC GAAACCCTCG GCAGCTCCGT GCAAAAGCAC       420

GCGGGCAAGG TGCTATTCGT GGCTATCCTG GTGCTGAGCA CCTTCTGCGT CGGCCTGAAG       480

AGCGCCCAGA TCCACTCCAA GGTGCACCAG CTGTGGATCC AGGAGGGCGG CCGGCTGGAG       540

GCGGAACTGG CCTACACACA GAAGACGATC GGCGAGGACG AGTCGGCCAC GCATCAGCTG       600

CTCATTCAGA CGACCCACGA CCCGAACGCC TCCGTCCTGC ATCCGCAGGC GCTGCTTGCC       660

CACCTGGAGG TCCTGGTCAA GGCCACCGCC GTCAAGGTGC ACCTCTACGA CACCGAATGG       720

GGGCTGCGCG ACATGTGCAA CATGCCGAGC ACGCCCTCCT TCGAGGGCAT CTACTACATC       780

GAGCAGATCC TGCGCCACCT CATTCCGTGC TCGATCATCA CGCCGCTGGA CTGTTTCTGG       840

GAGGGAAGCC AGCTGTTGGG TCCGGAATCA GCGGTCGTTA TACCAGGCCT CAACCAACGA       900

CTCCTGTGGA CCACCCTGAA TCCCGCCTCT GTGATGCAGT ATATGAAACA AAAGATGTCC       960

GAGGAAAAGA TCAGCTTCGA CTTCGAGACC GTGGAGCAGT ACATGAAGCG TGCGGCCATT      1020

GGCAGTGGCT ACATGGAGAA GCCCTGCCTG AACCCACTGA ATCCCAATTG CCCGGACACG      1080

GCACCGAACA AGAACAGCAC CCAGCCGCCG GATGTGGGAG CCATCCTGTC CGGAGGCTGC      1140

TACGGTTATG CCGCGAAGCA CATGCACTGG CCGGAGGAGC TGATTGTGGG CGGACGGAAG      1200

AGGAACCGCA GCGGACACTT GAGGAAGGCC CAGGCCCTGC AGTCGGTGGT GCAGCTGATG      1260

ACCGAGAAGG AAATGTACGA CCAGTGGCAG GACAACTACA AGGTGCACCA TCTTGGATGG      1320

ACGCAGGAGA AGGCAGCGGA GGTTTTGAAC GCCTGGCAGC GCAACTTTTC GCGGGAGGTG      1380

GAACAGCTGC TACGTAAACA GTCGAGAATT GCCACCAACT ACGATATCTA CGTGTTCAGC      1440
```

```
TCGGCTGCAC TGGATGACAT CCTGGCCAAG TTCTCCCATC CCAGCGCCTT GTCCATTGTC   1500
ATCGGCGTGG CCGTCACCGT TTTGTATGCC TTTTGCACGC TCCTCCGCTG GAGGGACCCC   1560
GTCCGTGGCC AGAGCAGTGT GGGCGTGGCC GGAGTTCTGC TCATGTGCTT CAGTACCGCC   1620
GCCGGATTGG GATTGTCAGC CCTGCTCGGT ATCGTTTTCA ATGCGCTGAC CGCTGCCTAT   1680
GCGGAGAGCA ATCGGCGGGA GCAGACCAAG CTGATTCTCA AGAACGCCAG CACCCAGGTG   1740
GTTCCGTTTT TGGCCCTTGG TCTGGGCGTC GATCACATCT TCATAGTGGG ACCGAGCATC   1800
CTGTTCAGTG CCTGCAGCAC CGCAGGATCC TTCTTTGCGG CCGCCTTTAT TCCGGTGCCG   1860
GCTTTGAAGG TATTCTGTCT GCAGGCTGCC ATCGTAATGT GCTCCAATTT GGCAGCGGCT   1920
CTATTGGTTT TTCCGGCCAT GATTTCGTTG GATCTACGGA GACGTACCGC CGGCAGGGCG   1980
GACATCTTCT GCTGCTGTTT TCCGGTGTGG AAGGAACAGC CGAAGGTGGC ACCTCCGGTG   2040
CTGCCGCTGA ACAACAACAA CGGGCGCGGG GCCCGGCATC CGAAGAGCTG CAACAACAAC   2100
AGGGTGCCGC TGCCCGCCCA GAATCCTCTG CTGGAACAGA GGGCAGACAT CCCTGGGAGC   2160
AGTCACTCAC TGGCGTCCTT CTCCCTGGCA ACCTTCGCCT TCAGCACTA CACTCCCTTC   2220
CTCATGCGCA GCTGGGTGAA GTTCCTGACC GTTATGGGTT TCCTGGCGGC CCTCATATCC   2280
AGCTTGTATG CCTCCACGCG CCTTCAGGAT GGCCTGGACA TTATTGATCT GGTGCCCAAG   2340
GACAGCAACG AGCACAAGTT CCTGGATGCT CAAACTCGGC TCTTTGGCTT CTACAGCATG   2400
TATGCGGTTA CCCAGGGCAA CTTTGAATAT CCCACCCAGC AGCAGTTGCT CAGGGACTAC   2460
CATGATTCCT TTGTGCGGGT GCCACATGTG ATCAAGAATG ATAACGGTGG ACTGCCGGAC   2520
TTCTGGCTGC TGCTCTTCAG CGAGTGGCTG GGTAATCTGC AAAAGATATT CGACGAGGAA   2580
TACCGCGACG GACGGCTGAC CAAGGAGTGC TGGTTCCCAA ACGCCAGCAG CGATGCCATC   2640
CTGGCCTACA AGCTAATCGT GCAAACCGGC CATGTGGACA ACCCCGTGGA CAAGGAACTG   2700
GTGCTCACCA ATCGCCTGGT CAACAGCGAT GGCATCATCA ACCAACGCGC CTTCTACAAC   2760
TATCTGTCGG CATGGGCCAC CAACGACGTC TTCGCCTACG GAGCTTCTCA GGGCAAATTG   2820
TATCCGGAAC CGCGCCAGTA TTTTCACCAA CCCAACGAGT ACGATCTTAA GATACCCAAG   2880
AGTCTGCCAT TGGTCTACGC TCAGATGCCC TTTTACCTCC ACGGACTAAC AGATACCTCG   2940
CAGATCAAGA CCCTGATAGG TCATATTCGC GACCTGAGCG TCAAGTACGA GGGCTTCGGC   3000
CTGCCCAACT ATCCATCGGG CATTCCCTTC ATCTTCTGGG AGCAGTACAT GACCCTGCGC   3060
TCCTCACTGG CCATGATCCT GGCCTGCGTG CTACTCGCCG CCCTGGTGCT GGTCTCCCTG   3120
CTCCTGCTCT CCGTTTGGGC CGCCGTTCTC GTGATCCTCA GCGTTCTGGC CTCGCTGGCC   3180
CAGATCTTTG GGGCCATGAC TCTGCTGGGC ATCAAACTCT CGGCCATTCC GGCAGTCATA   3240
CTCATCCTCA GCGTGGGCAT GATGCTGTGC TTCAATGTGC TGATATCACT GGGCTTCATG   3300
ACATCCGTTG GCAACCGACA GCGCCGCGTC CAGCTGAGCA TGCAGATGTC CCTGGGACCA   3360
```

```
CTTGTCCACG GCATGCTGAC CTCCGGAGTG GCCGTGTTCA TGCTCTCCAC GTCGCCCTTT    3420

GAGTTTGTGA TCCGGCACTT CTGCTGGCTT CTGCTGGTGG TCTTATGCGT TGGCGCCTGC    3480

AACAGCCTTT TGGTGTTCCC CATCCTACTG AGCATGGTGG GACCGGAGGC GGAGCTGGTG    3540

CCGCTGGAGC ATCCAGACCG CATATCCACG CCCTCTCCGC TGCCCGTGCG CAGCAGCAAG    3600

AGATCGGGCA AATCCTATGT GGTGCAGGGA TCGCGATCCT CGCGAGGCAG CTGCCAGAAG    3660

TCGCATCACC ACCACCACAA AGACCTTAAT GATCCATCGC TGACGACGAT CACCGAGGAG    3720

CCGCAGTCGT GGAAGTCCAG CAACTCGTCC ATCCAGATGC CCAATGATTG GACCTACCAG    3780

CCGCGGGAAC AGCGACCCGC CTCCTACGCG GCCCCGCCCC CCGCCTATCA CAAGGCCGCC    3840

GCCCAGCAGC ACCACCAGCA TCAGGGCCCG CCCACAACGC CCCCGCCTCC CTTCCCGACG    3900

GCCTATCCGC CGGAGCTGCA GAGCATCGTG GTGCAGCCGG AGGTGACGGT GGAGACGACG    3960

CACTCGGACA GCAACACCAC CAAGGTGACG GCCACGGCCA ACATCAAGGT GGAGCTGGCC    4020

ATGCCCGGCA GGGCGGTGCG CAGCTATAAC TTTACGAGTT AGCACTAGCA CTAGTTCCTG    4080

TAGCTATTAG GACGTATCTT TAGACTCTAG CCTAAGCCGT AACCCTATTT GTATCTGTAA    4140

AATCGATTTG TCCAGCGGGT CTGCTGAGGA TTTCGTTCTC ATGGATTCTC ATGGATTCTC    4200

ATGGATGCTT AAATGGCATG GTAATTGGCA AAATATCAAT TTTTGTGTCT CAAAAAGATG    4260

CATTAGCTTA TGGTTTCAAG ATACATTTTT AAAGAGTCCG CCAGATATTT ATATAAAAAA    4320

AATCCAAAAT CGACGTATCC ATGAAAATTG AAAAGCTAAG CAGACCCGTA TGTATGTATA    4380

TGTGTATGCA TGTTAGTTAA TTTCCCGAAG TCCGGTATTT ATAGCAGCTG CCTT           4434
```

(2) INFORMATION FOR SEQ ID NO:6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1285 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Asp Arg Asp Ser Leu Pro Arg Val Pro Asp Thr His Gly Asp Val
1               5                   10                  15

Val Asp Glu Lys Leu Phe Ser Asp Leu Tyr Ile Arg Thr Ser Trp Val
            20                  25                  30

Asp Ala Gln Val Ala Leu Asp Gln Ile Asp Lys Gly Lys Ala Arg Gly
        35                  40                  45

Ser Arg Thr Ala Ile Tyr Leu Arg Ser Val Phe Gln Ser His Leu Glu
    50                  55                  60
```

```
Thr Leu Gly Ser Ser Val Gln Lys His Ala Gly Lys Val Leu Phe Val
65                  70                  75                  80

Ala Ile Leu Val Leu Ser Thr Phe Cys Val Gly Leu Lys Ser Ala Gln
                85                  90                  95

Ile His Ser Lys Val His Gln Leu Trp Ile Gln Glu Gly Gly Arg Leu
            100                 105                 110

Glu Ala Glu Leu Ala Tyr Thr Gln Lys Thr Ile Gly Glu Asp Glu Ser
            115                 120                 125

Ala Thr His Gln Leu Leu Ile Gln Thr Thr His Asp Pro Asn Ala Ser
        130                 135                 140

Val Leu His Pro Gln Ala Leu Leu Ala His Leu Glu Val Leu Val Lys
145                 150                 155                 160

Ala Thr Ala Val Lys Val His Leu Tyr Asp Thr Glu Trp Gly Leu Arg
                165                 170                 175

Asp Met Cys Asn Met Pro Ser Thr Pro Ser Phe Glu Gly Ile Tyr Tyr
            180                 185                 190

Ile Glu Gln Ile Leu Arg His Leu Ile Pro Cys Ser Ile Ile Thr Pro
            195                 200                 205

Leu Asp Cys Phe Trp Glu Gly Ser Gln Leu Leu Gly Pro Glu Ser Ala
210                 215                 220

Val Val Ile Pro Gly Leu Asn Gln Arg Leu Leu Trp Thr Thr Leu Asn
225                 230                 235                 240

Pro Ala Ser Val Met Gln Tyr Met Lys Gln Lys Met Ser Glu Glu Lys
                245                 250                 255

Ile Ser Phe Asp Phe Glu Thr Val Glu Gln Tyr Met Lys Arg Ala Ala
            260                 265                 270

Ile Gly Ser Gly Tyr Met Glu Lys Pro Cys Leu Asn Pro Leu Asn Pro
            275                 280                 285

Asn Cys Pro Asp Thr Ala Pro Asn Lys Asn Ser Thr Gln Pro Pro Asp
        290                 295                 300

Val Gly Ala Ile Leu Ser Gly Gly Cys Tyr Gly Tyr Ala Ala Lys His
305                 310                 315                 320

Met His Trp Pro Glu Glu Leu Ile Val Gly Gly Arg Lys Arg Asn Arg
                325                 330                 335

Ser Gly His Leu Arg Lys Ala Gln Ala Leu Gln Ser Val Val Gln Leu
            340                 345                 350

Met Thr Glu Lys Glu Met Tyr Asp Gln Trp Gln Asp Asn Tyr Lys Val
            355                 360                 365

His His Leu Gly Trp Thr Gln Glu Lys Ala Ala Glu Val Leu Asn Ala
        370                 375                 380

Trp Gln Arg Asn Phe Ser Arg Glu Val Glu Gln Leu Leu Arg Lys Gln
385                 390                 395                 400
```

Ser Arg Ile Ala Thr Asn Tyr Asp Ile Tyr Val Phe Ser Ser Ala Ala
           405   410   415

Leu Asp Asp Ile Leu Ala Lys Phe Ser His Pro Ser Ala Leu Ser Ile
           420   425   430

Val Ile Gly Val Ala Val Thr Val Leu Tyr Ala Phe Cys Thr Leu Leu
           435   440   445

Arg Trp Arg Asp Pro Val Arg Gly Gln Ser Ser Val Gly Val Ala Gly
           450   455   460

Val Leu Leu Met Cys Phe Ser Thr Ala Ala Gly Leu Gly Leu Ser Ala
465   470   475   480

Leu Leu Gly Ile Val Phe Asn Ala Leu Thr Ala Ala Tyr Ala Glu Ser
           485   490   495

Asn Arg Arg Glu Gln Thr Lys Leu Ile Leu Lys Asn Ala Ser Thr Gln
           500   505   510

Val Val Pro Phe Leu Ala Leu Gly Leu Gly Val Asp His Ile Phe Ile
           515   520   525

Val Gly Pro Ser Ile Leu Phe Ser Ala Cys Ser Thr Ala Gly Ser Phe
           530   535   540

Phe Ala Ala Ala Phe Ile Pro Val Pro Ala Leu Lys Val Phe Cys Leu
545   550   555   560

Gln Ala Ala Ile Val Met Cys Ser Asn Leu Ala Ala Ala Leu Leu Val
           565   570   575

Phe Pro Ala Met Ile Ser Leu Asp Leu Arg Arg Arg Thr Ala Gly Arg
           580   585   590

Ala Asp Ile Phe Cys Cys Cys Phe Pro Val Trp Lys Glu Gln Pro Lys
           595   600   605

Val Ala Pro Pro Val Leu Pro Leu Asn Asn Asn Asn Gly Arg Gly Ala
           610   615   620

Arg His Pro Lys Ser Cys Asn Asn Asn Arg Val Pro Leu Pro Ala Gln
625   630   635   640

Asn Pro Leu Leu Glu Gln Arg Ala Asp Ile Pro Gly Ser Ser His Ser
           645   650   655

Leu Ala Ser Phe Ser Leu Ala Thr Phe Ala Phe Gln His Tyr Thr Pro
           660   665   670

Phe Leu Met Arg Ser Trp Val Lys Phe Leu Thr Val Met Gly Phe Leu
           675   680   685

Ala Ala Leu Ile Ser Ser Leu Tyr Ala Ser Thr Arg Leu Gln Asp Gly
           690   695   700

Leu Asp Ile Ile Asp Leu Val Pro Lys Asp Ser Asn Glu His Lys Phe
           705   710   715   720

Leu Asp Ala Gln Thr Arg Leu Phe Gly Phe Tyr Ser Met Tyr Ala Val
           725   730   735

Thr Gln Gly Asn Phe Glu Tyr Pro Thr Gln Gln Gln Leu Leu Arg Asp

```
                740                    745                      750
      Tyr His Asp Ser Phe Arg Val Pro His Val Ile Lys Asn Asp Asn Gly
              755                 760                 765
      Gly Leu Pro Asp Phe Trp Leu Leu Leu Phe Ser Glu Trp Leu Gly Asn
              770                 775                 780
      Leu Gln Lys Ile Phe Asp Glu Glu Tyr Arg Asp Gly Arg Leu Thr Lys
      785                 790                 795                 800
      Glu Cys Trp Phe Pro Asn Ala Ser Ser Asp Ala Ile Leu Ala Tyr Lys
                      805                 810                 815
      Leu Ile Val Gln Thr Gly His Val Asp Asn Pro Val Asp Lys Glu Leu
                  820                 825                 830
      Val Leu Thr Asn Arg Leu Val Asn Ser Asp Gly Ile Ile Asn Gln Arg
          835                 840                 845
      Ala Phe Tyr Asn Tyr Leu Ser Ala Trp Ala Thr Asn Asp Val Phe Ala
          850                 855                 860
      Tyr Gly Ala Ser Gln Gly Lys Leu Tyr Pro Glu Pro Arg Gln Tyr Phe
      865                 870                 875                 880
      His Gln Pro Asn Glu Tyr Asp Leu Lys Ile Pro Lys Ser Leu Pro Leu
                      885                 890                 895
      Val Tyr Ala Gln Met Pro Phe Tyr Leu His Gly Leu Thr Asp Thr Ser
                  900                 905                 910
      Gln Ile Lys Thr Leu Ile Gly His Ile Arg Asp Leu Ser Val Lys Tyr
              915                 920                 925
      Glu Gly Phe Gly Leu Pro Asn Tyr Pro Ser Gly Ile Pro Phe Ile Phe
          930                 935                 940
      Trp Glu Gln Tyr Met Thr Leu Arg Ser Ser Leu Ala Met Ile Leu Ala
      945                 950                 955                 960
      Cys Val Leu Leu Ala Ala Leu Val Leu Val Ser Leu Leu Leu Ser
                      965                 970                 975
      Val Trp Ala Ala Val Leu Val Ile Leu Ser Val Leu Ala Ser Leu Ala
                  980                 985                 990
      Gln Ile Phe Gly Ala Met Thr Leu Leu Gly Ile Lys Leu Ser Ala Ile
              995                 1000                1005
      Pro Ala Val Ile Leu Ile Leu Ser Val Gly Met Met Leu Cys Phe Asn
          1010                1015                1020
      Val Leu Ile Ser Leu Gly Phe Met Thr Ser Val Gly Asn Arg Gln Arg
      1025                1030                1035                1040
      Arg Val Gln Leu Ser Met Gln Met Ser Leu Gly Pro Leu Val His Gly
                  1045                1050                1055
      Met Leu Thr Ser Gly Val Ala Val Phe Met Leu Ser Thr Ser Pro Phe
              1060                1065                1070
      Glu Phe Val Ile Arg His Phe Cys Trp Leu Leu Leu Val Val Leu Cys
```

```
            1075                1080                1085
     Val Gly Ala Cys Asn Ser Leu Leu Val Phe Pro Ile Leu Leu Ser Met
         1090                1095                1100

     Val Gly Pro Glu Ala Glu Leu Val Pro Leu Glu His Pro Asp Arg Ile
     1105                1110                1115                1120

     Ser Thr Pro Ser Pro Leu Pro Val Arg Ser Ser Lys Arg Ser Gly Lys
                     1125                1130                1135

     Ser Tyr Val Val Gln Gly Ser Arg Ser Ser Arg Gly Ser Cys Gln Lys
                 1140                1145                1150

     Ser His His His His His Lys Asp Leu Asn Asp Pro Ser Leu Thr Thr
             1155                1160                1165

     Ile Thr Glu Glu Pro Gln Ser Trp Lys Ser Ser Asn Ser Ser Ile Gln
             1170                1175                1180

     Met Pro Asn Asp Trp Thr Tyr Gln Pro Arg Glu Gln Arg Pro Ala Ser
     1185                1190                1195                1200

     Tyr Ala Ala Pro Pro Pro Ala Tyr His Lys Ala Ala Ala Gln Gln His
                     1205                1210                1215

     His Gln His Gln Gly Pro Pro Thr Thr Pro Pro Pro Pro Phe Pro Thr
                 1220                1225                1230

     Ala Tyr Pro Pro Glu Leu Gln Ser Ile Val Val Gln Pro Glu Val Thr
                 1235                1240                1245

     Val Glu Thr Thr His Ser Asp Ser Asn Thr Thr Lys Val Thr Ala Thr
         1250                1255                1260

     Ala Asn Ile Lys Val Glu Leu Ala Met Pro Gly Arg Ala Val Arg Ser
     1265                1270                1275                1280

     Tyr Asn Phe Thr Ser
                 1285
```

(2) INFORMATION FOR SEQ ID NO:7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 345 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
AAGGTCCATC AGCTTTGGAT ACAGGAAGGT GGTTCGCTCG AGCATGAGCT AGCCTACACG      60

CAGAAATCGC TCGGCGAGAT GGACTCCTCC ACGCACCAGC TGCTAATCCA AACNCCCAAA     120

GATATGGACG CCTCGATACT GCACCCGAAC GCGCTACTGA CGCACCTGGA CGTGGTGAAG     180

AAAGCGATCT CGGTGACGGT GCACATGTAC GACATCACGT GGAGNCTCAA GGACATGTGC     240
```

TACTCGCCCA GCATACCGAG NTTCGATACG CACTTTATCG AGCAGATCTT CGAGAACATC   300

ATACCGTGCG CGATCATCAC GCCGCTGGAT TGCTTTTGGG AGGGA   345

(2) INFORMATION FOR SEQ ID NO:8:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 115 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Lys Val His Gln Leu Trp Ile Gln Glu Gly Gly Ser Leu Glu His Glu
1   5   10   15

Leu Ala Tyr Thr Gln Lys Ser Leu Gly Glu Met Asp Ser Ser Thr His
            20   25   30

Gln Leu Leu Ile Gln Thr Pro Lys Asp Met Asp Ala Ser Ile Leu His
        35   40   45

Pro Asn Ala Leu Leu Thr His Leu Asp Val Val Lys Lys Ala Ile Ser
    50   55   60

Val Thr Val His Met Tyr Asp Ile Thr Trp Xaa Leu Lys Asp Met Cys
65   70   75   80

Tyr Ser Pro Ser Ile Pro Xaa Phe Asp Thr His Phe Ile Glu Gln Ile
                85   90   95

Phe Glu Asn Ile Ile Pro Cys Ala Ile Ile Thr Pro Leu Asp Cys Phe
        100   105   110

Trp Glu Gly
    115

(2) INFORMATION FOR SEQ ID NO:9:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5187 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGGTCTGTCA CCCGGAGCCG GAGTCCCCGG CGGCCAGCAG CGTCCTCGCG AGCCGAGCGC   60

CCAGGCGCGC CCGGAGCCCG CGGCGGCGGC GGCAACATGG CCTCGGCTGG TAACGCCGCC   120

```
GGGGCCCTGG GCAGGCAGGC CGGCGGCGGG AGGCGCAGAC GGACCGGGGG ACCGCACCGC      180
GCCGCGCCGG ACCGGGACTA TCTGCACCGG CCCAGCTACT GCGACGCCGC CTTCGCTCTG      240
GAGCAGATTT CCAAGGGGAA GGCTACTGGC CGGAAAGCGC CGCTGTGGCT GAGAGCGAAG      300
TTTCAGAGAC TCTTATTTAA ACTGGGTTGT TACATTCAAA AGAACTGCGG CAAGTTTTTG      360
GTTGTGGGTC TCCTCATATT TGGGGCCTTC GCTGTGGGAT TAAAGGCAGC TAATCTCGAG      420
ACCAACGTGG AGGAGCTGTG GGTGGAAGTT GGTGGACGAG TGAGTCGAGA ATTAAATTAT      480
ACCCGTCAGA AGATAGGAGA AGAGGCTATG TTTAATCCTC AACTCATGAT ACAGACTCCA      540
AAAGAAGAAG GCGCTAATGT TCTGACCACA GAGGCTCTCC TGCAACACCT GGACTCAGCA      600
CTCCAGGCCA GTCGTGTGCA CGTCTACATG TATAACAGGC AATGGAAGTT GGAACATTTG      660
TGCTACAAAT CAGGGGAACT TATCACGGAG ACAGGTTACA TGGATCAGAT AATAGAATAC      720
CTTTACCCTT GCTTAATCAT TACACCTTTG GACTGCTTCT GGGAAGGGGC AAAGCTACAG      780
TCCGGGACAG CATACCTCCT AGGTAAGCCT CCTTTACGGT GGACAAACTT GACCCCTTG      840
GAATTCCTAG AAGAGTTAAA GAAAATAAAC TACCAAGTGG ACAGCTGGGA GGAAATGCTG      900
AATAAAGCCG AAGTTGGCCA TGGGTACATG GACCGGCCTT GCCTCAACCC AGCCGACCCA      960
GATTGCCCTG CCACAGCCCC TAACAAAAAT TCAACCAAAC CTCTTGATGT GGCCCTTGTT     1020
TTGAATGGTG GATGTCAAGG TTTATCCAGG AAGTATATGC ATTGGCAGGA GGAGTTGATT     1080
GTGGGTGGTA CCGTCAAGAA TGCCACTGGA AAACTTGTCA GCGCTCACGC CCTGCAAACC     1140
ATGTTCCAGT TAATGACTCC CAAGCAAATG TATGAACACT TCAGGGGCTA CGACTATGTC     1200
TCTCACATCA ACTGGAATGA AGACAGGGCA GCCGCCATCC TGGAGGCCTG GCAGAGGACT     1260
TACGTGGAGG TGGTTCATCA AAGTGTCGCC CCAAACTCCA CTCAAAAGGT GCTTCCCTTC     1320
ACAACCACGA CCCTGGACGA CATCCTAAAA TCCTTCTCTG ATGTCAGTGT CATCCGAGTG     1380
GCCAGCGGCT ACCTACTGAT GCTTGCCTAT GCCTGTTTAA CCATGCTGCG CTGGGACTGC     1440
TCCAAGTCCC AGGGTGCCGT GGGGCTGGCT GGCGTCCTGT TGGTTGCGCT GTCAGTGGCT     1500
GCAGGATTGG GCCTCTGCTC CTTGATTGGC ATTTCTTTTA ATGCTGCGAC AACTCAGGTT     1560
TTGCCGTTTC TTGCTCTTGG TGTTGGTGTG GATGATGTCT TCCTCCTGGC CCATGCATTC     1620
AGTGAAACAG GACAGAATAA GAGGATTCCA TTTGAGGACA GGACTGGGGA GTGCCTCAAG     1680
CGCACCGGAG CCAGCGTGGC CCTCACCTCC ATCAGCAATG TCACCGCCTT CTTCATGGCC     1740
GCATTGATCC CTATCCCTGC CCTGCGAGCG TTCTCCCTCC AGGCTGCTGT GGTGGTGGTA     1800
TTCAATTTTG CTATGGTTCT GCTCATTTTT CCTGCAATTC TCAGCATGGA TTTATACAGA     1860
CGTGAGGACA GAAGATTGGA TATTTTCTGC TGTTTCACAA GCCCCTGTGT CAGCAGGGTG     1920
ATTCAAGTTG AGCCACAGGC CTACACAGAG CCTCACAGTA ACACCCGGTA CAGCCCCCCA     1980
CCCCCATACA CCAGCCACAG CTTCGCCCAC GAAACCCATA TCACTATGCA GTCCACCGTT     2040
```

CAGCTCCGCA CAGAGTATGA CCCTCACACG CACGTGTACT ACACCACCGC CGAGCCACGC   2100

TCTGAGATCT CTGTACAGCC TGTTACCGTC ACCCAGGACA ACCTCAGCTG TCAGAGTCCC   2160

GAGAGCACCA GCTCTACCAG GGACCTGCTC TCCCAGTTCT CAGACTCCAG CCTCCACTGC   2220

CTCGAGCCCC CCTGCACCAA GTGGACACTC TCTTCGTTTG CAGAGAAGCA CTATGCTCCT   2280

TTCCTCCTGA AACCCAAAGC CAAGGTTGTG GTAATCCTTC TTTTCCTGGG CTTGCTGGGG   2340

GTCAGCCTTT ATGGGACCAC CCGAGTGAGA GACGGGCTGG ACCTCACGGA CATTGTTCCC   2400

CGGGAAACCA GAGAATATGA CTTCATAGCT GCCCAGTTCA AGTACTTCTC TTTCTACAAC   2460

ATGTATATAG TCACCCAGAA AGCAGACTAC CCGAATATCC AGCACCTACT TTACGACCTT   2520

CATAAGAGTT TCAGCAATGT GAAGTATGTC ATGCTGGAGG AGAACAAGCA ACTTCCCCAA   2580

ATGTGGCTGC ACTACTTTAG AGACTGGCTT CAAGGACTTC AGGATGCATT TGACAGTGAC   2640

TGGGAAACTG GGAGGATCAT GCCAAACAAT TATAAAAATG GATCAGATGA CGGGGTCCTC   2700

GCTTACAAAC TCCTGGTGCA GACTGGCAGC CGAGACAAGC CCATCGACAT TAGTCAGTTG   2760

ACTAAACAGC GTCTGGTAGA CGCAGATGGC ATCATTAATC CGAGCGCTTT CTACATCTAC   2820

CTGACCGCTT GGGTCAGCAA CGACCCTGTA GCTTACGCTG CCTCCCAGGC CAACATCCGG   2880

CCTCACCGGC CGGAGTGGGT CCATGACAAA GCCGACTACA TGCCAGAGAC CAGGCTGAGA   2940

ATCCCAGCAG CAGAGCCCAT CGAGTACGCT CAGTTCCCTT TCTACCTCAA CGGCCTACGA   3000

GACACCTCAG ACTTTGTGGA AGCCATAGAA AAAGTGAGAG TCATCTGTAA CAACTATACG   3060

AGCCTGGGAC TGTCCAGCTA CCCCAATGGC TACCCCTTCC TGTTCTGGGA GCAATACATC   3120

AGCCTGCGCC ACTGGCTGCT GCTATCCATC AGCGTGGTGC TGGCCTGCAC GTTTCTAGTG   3180

TGCGCAGTCT TCCTCCTGAA CCCCTGGACG GCCGGGATCA TTGTCATGGT CCTGGCTCTG   3240

ATGACCGTTG AGCTCTTTGG CATGATGGGC CTCATTGGGA TCAAGCTGAG TGCTGTGCCT   3300

GTGGTCATCC TGATTGCATC TGTTGGCATC GGAGTGGAGT TCACCGTCCA CGTGGCTTTG   3360

GCCTTTCTGA CAGCCATTGG GGACAAGAAC CACAGGGCTA TGCTCGCTCT GGAACACATG   3420

TTTGCTCCCG TTCTGGACGG TGCTGTGTCC ACTCTGCTGG GTGTACTGAT GCTTGCAGGG   3480

TCCGAATTTG ATTTCATTGT CAGATACTTC TTTGCCGTCC TGGCCATTCT CACCGTCTTG   3540

GGGGTTCTCA ATGGACTGGT TCTGCTGCCT GTCCTCTTAT CCTTCTTTGG ACCGTGTCCT   3600

GAGGTGTCTC CAGCCAATGG CCTAAACCGA CTGCCCACTC CTTCGCCTGA GCCGCCTCCA   3660

AGTGTCGTCC GGTTTGCCGT GCCTCCTGGT CACACGAACA ATGGGTCTGA TTCCTCCGAC   3720

TCGGAGTACA GCTCTCAGAC CACGGTGTCT GGCATCAGTG AGGAGCTCAG GCAATACGAA   3780

GCACAGCAGG GTGCCGGAGG CCCTGCCCAC CAAGTGATTG TGGAAGCCAC AGAAAACCCT   3840

GTCTTTGCCC GGTCCACTGT GGTCCATCCG GACTCCAGAC ATCAGCCTCC CTTGACCCCT   3900

CGGCAACAGC CCCACCTGGA CTCTGGCTCC TTGTCCCCTG GACGGCAAGG CCAGCAGCCT   3960

CGAAGGGATC CCCCTAGAGA AGGCTTGCGG CCACCCCCCT ACAGACCGCG CAGAGACGCT   4020

TTTGAAATTT CTACTGAAGG GCATTCTGGC CCTAGCAATA GGGACCGCTC AGGGCCCCGT   4080

GGGGCCCGTT CTCACAACCC TCGGAACCCA ACGTCCACCG·CCATGGGCAG CTCTGTGCCC   4140

AGCTACTGCC AGCCCATCAC CACTGTGACG GCTTCTGCTT CGGTGACTGT TGCTGTGCAT   4200

CCCCCGCCTG GACCTGGGCG CAACCCCCGA GGGGGGCCCT GTCCAGGCTA TGAGAGCTAC   4260

CCTGAGACTG ATCACGGGGT ATTTGAGGAT CCTCATGTGC CTTTTCATGT CAGGTGTGAG   4320

AGGAGGGACT CAAAGGTGGA GGTCATAGAG CTACAGGACG TGGAATGTGA GGAGAGGCCG   4380

TGGGGGAGCA GCTCCAACTG AGGGTAATTA AAATCTGAAG CAAAGAGGCC AAAGATTGGA   4440

AAGCCCCGCC CCCACCTCTT TCCAGAACTG CTTGAAGAGA ACTGCTTGGA ATTATGGGAA   4500

GGCAGTTCAT TGTTACTGTA ACTGATTGTA TTATTKKGTG AAATATTTCT ATAAATATTT   4560

AARAGGTGTA CACATGTAAT ATACATGGAA ATGCTGTACA GTCTATTTCC TGGGGCCTCT   4620

CCACTCCTGC CCCAGAGTGG GGAGACCACA GGGGCCCTTT CCCCTGTGTA CATTGGTCTC   4680

TGTGCCACAA CCAAGCTTAA CTTAGTTTTA AAAAAAATCT CCCAGCATAT GTCGCTGCTG   4740

CTTAAATATT GTATAATTTA CTTGTATAAT TCTATGCAAA TATTGCTTAT GTAATAGGAT   4800

TATTTGTAAA GGTTTCTGTT TAAAATATTT TAAATTTGCA TATCACAACC CTGTGGTAGG   4860

ATGAATTGTT ACTGTTAACT TTTGAACACG CTATGCGTGG TAATTGTTTA ACGAGCAGAC   4920

ATGAAGAAAA CAGGTTAATC CCAGTGGCTT CTCTAGGGGT AGTTGTATAT GGTTCGCATG   4980

GGTGGATGTG TGTGTGCATG TGACTTTCCA ATGTACTGTA TTGTGGTTTG TTGTTGTTGT   5040

TGCTGTTGTT GTTCATTTTG GTGTTTTTGG TTGCTTTGTA TGATCTTAGC TCTGGCCTAG   5100

GTGGGCTGGG AAGGTCCAGG TCTTTTTCTG TCGTGATGCT GGTGGAAAGG TGACCCCAAT   5160

CATCTGTCCT ATTCTCTGGG ACTATTC   5187

(2) INFORMATION FOR SEQ ID NO:10:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 1434 amino acids
          (B) TYPE: amino acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: protein

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ala Ser Ala Gly Asn Ala Ala Gly Ala Leu Gly Arg Gln Ala Gly
1   5   10   15

Gly Gly Arg Arg Arg Arg Thr Gly Gly Pro His Arg Ala Ala Pro Asp

```
                    20              25              30
Arg Asp Tyr Leu His Arg Pro Ser Tyr Cys Asp Ala Ala Phe Ala Leu
        35              40              45

Glu Gln Ile Ser Lys Gly Lys Ala Thr Gly Arg Lys Ala Pro Leu Trp
        50              55              60

Leu Arg Ala Lys Phe Gln Arg Leu Leu Phe Lys Leu Gly Cys Tyr Ile
65              70              75              80

Gln Lys Asn Cys Gly Lys Phe Leu Val Val Gly Leu Leu Ile Phe Gly
                85              90              95

Ala Phe Ala Val Gly Leu Lys Ala Ala Asn Leu Glu Thr Asn Val Glu
            100             105             110

Glu Leu Trp Val Glu Val Gly Gly Arg Val Ser Arg Glu Leu Asn Tyr
            115             120             125

Thr Arg Gln Lys Ile Gly Glu Glu Ala Met Phe Asn Pro Gln Leu Met
        130             135             140

Ile Gln Thr Pro Lys Glu Glu Gly Ala Asn Val Leu Thr Thr Glu Ala
        145             150             155             160

Leu Leu Gln His Leu Asp Ser Ala Leu Gln Ala Ser Arg Val His Val
                165             170             175

Tyr Met Tyr Asn Arg Gln Trp Lys Leu Glu His Leu Cys Tyr Lys Ser
            180             185             190

Gly Glu Leu Ile Thr Glu Thr Gly Tyr Met Asp Gln Ile Ile Glu Tyr
        195             200             205

Leu Tyr Pro Cys Leu Ile Ile Thr Pro Leu Asp Cys Phe Trp Glu Gly
    210             215             220

Ala Lys Leu Gln Ser Gly Thr Ala Tyr Leu Leu Gly Lys Pro Pro Leu
225             230             235             240

Arg Trp Thr Asn Phe Asp Pro Leu Glu Phe Leu Glu Glu Leu Lys Lys
                245             250             255

Ile Asn Tyr Gln Val Asp Ser Trp Glu Glu Met Leu Asn Lys Ala Glu
            260             265             270

Val Gly His Gly Tyr Met Asp Arg Pro Cys Leu Asn Pro Ala Asp Pro
        275             280             285

Asp Cys Pro Ala Thr Ala Pro Asn Lys Asn Ser Thr Lys Pro Leu Asp
    290             295             300

Val Ala Leu Val Leu Asn Gly Gly Cys Gln Gly Leu Ser Arg Lys Tyr
305             310             315             320

Met His Trp Gln Glu Glu Leu Ile Val Gly Gly Thr Val Lys Asn Ala
            325             330             335

Thr Gly Lys Leu Val Ser Ala His Ala Leu Gln Thr Met Phe Gln Leu
        340             345             350

Met Thr Pro Lys Gln Met Tyr Glu His Phe Arg Gly Tyr Asp Tyr Val
    355             360             365
```

Ser His Ile Asn Trp Asn Glu Asp Arg Ala Ala Ala Ile Leu Glu Ala
370   375   380

Trp Gln Arg Thr Tyr Val Glu Val Val His Gln Ser Val Ala Pro Asn
385   390   395   400

Ser Thr Gln Lys Val Leu Pro Phe Thr Thr Thr Thr Leu Asp Asp Ile
           405   410   415

Leu Lys Ser Phe Ser Asp Val Ser Val Ile Arg Val Ala Ser Gly Tyr
           420   425   430

Leu Leu Met Leu Ala Tyr Ala Cys Leu Thr Met Leu Arg Trp Asp Cys
           435   440   445

Ser Lys Ser Gln Gly Ala Val Gly Leu Ala Gly Val Leu Leu Val Ala
           450   455   460

Leu Ser Val Ala Ala Gly Leu Gly Leu Cys Ser Leu Ile Gly Ile Ser
465   470   475   480

Phe Asn Ala Ala Thr Thr Gln Val Leu Pro Phe Leu Ala Leu Gly Val
           485   490   495

Gly Val Asp Asp Val Phe Leu Leu Ala His Ala Phe Ser Glu Thr Gly
           500   505   510

Gln Asn Lys Arg Ile Pro Phe Glu Asp Arg Thr Gly Glu Cys Leu Lys
           515   520   525

Arg Thr Gly Ala Ser Val Ala Leu Thr Ser Ile Ser Asn Val Thr Ala
         530   535   540

Phe Phe Met Ala Ala Leu Ile Pro Ile Pro Ala Leu Arg Ala Phe Ser
545   550   555   560

Leu Gln Ala Ala Val Val Val Val Phe Asn Phe Ala Met Val Leu Leu
           565   570   575

Ile Phe Pro Ala Ile Leu Ser Met Asp Leu Tyr Arg Arg Glu Asp Arg
         580   585   590

Arg Leu Asp Ile Phe Cys Cys Phe Thr Ser Pro Cys Val Ser Arg Val
         595   600   605

Ile Gln Val Glu Pro Gln Ala Tyr Thr Glu Pro His Ser Asn Thr Arg
    610   615   620

Tyr Ser Pro Pro Pro Pro Tyr Thr Ser His Ser Phe Ala His Glu Thr
625   630   635   640

His Ile Thr Met Gln Ser Thr Val Gln Leu Arg Thr Glu Tyr Asp Pro
           645   650   655

His Thr His Val Tyr Tyr Thr Thr Ala Glu Pro Arg Ser Glu Ile Ser
         660   665   670

Val Gln Pro Val Thr Val Thr Gln Asp Asn Leu Ser Cys Gln Ser Pro
         675   680   685

Glu Ser Thr Ser Ser Thr Arg Asp Leu Leu Ser Gln Phe Ser Asp Ser
    690   695   700

Ser Leu His Cys Leu Glu Pro Pro Cys Thr Lys Trp Thr Leu Ser Ser
705   710   715   720

Phe Ala Glu Lys His Tyr Ala Pro Phe Leu Leu Lys Pro Lys Ala Lys
             725   730   735

Val Val Val Ile Leu Leu Phe Leu Gly Leu Leu Gly Val Ser Leu Tyr
          740   745   750
Gly Thr Thr Arg Val Arg Asp Gly Leu Asp Leu Thr Asp Ile Val Pro
          755   760   765

Arg Glu Thr Arg Glu Tyr Asp Phe Ile Ala Ala Gln Phe Lys Tyr Phe
       770   775   780

Ser Phe Tyr Asn Met Tyr Ile Val Thr Gln Lys Ala Asp Tyr Pro Asn
785   790   795   800

Ile Gln His Leu Leu Tyr Asp Leu His Lys Ser Phe Ser Asn Val Lys
             805   810   815

Tyr Val Met Leu Glu Glu Asn Lys Gln Leu Pro Gln Met Trp Leu His
          820   825   830

Tyr Phe Arg Asp Trp Leu Gln Gly Leu Gln Asp Ala Phe Asp Ser Asp
          835   840   845

Trp Glu Thr Gly Arg Ile Met Pro Asn Asn Tyr Lys Asn Gly Ser Asp
          850   855   860

Asp Gly Val Leu Ala Tyr Lys Leu Leu Val Gln Thr Gly Ser Arg Asp
865   870   875   880

Lys Pro Ile Asp Ile Ser Gln Leu Thr Lys Gln Arg Leu Val Asp Ala
             885   890   895

Asp Gly Ile Ile Asn Pro Ser Ala Phe Tyr Ile Tyr Leu Thr Ala Trp
          900   905   910

Val Ser Asn Asp Pro Val Ala Tyr Ala Ala Ser Gln Ala Asn Ile Arg
          915   920   925

Pro His Arg Pro Glu Trp Val His Asp Lys Ala Asp Tyr Met Pro Glu
          930   935   940

Thr Arg Leu Arg Ile Pro Ala Ala Glu Pro Ile Glu Tyr Ala Gln Phe
945   950   955   960

Pro Phe Tyr Leu Asn Gly Leu Arg Asp Thr Ser Asp Phe Val Glu Ala
             965   970   975

Ile Glu Lys Val Arg Val Ile Cys Asn Asn Tyr Thr Ser Leu Gly Leu
          980   985   990

Ser Ser Tyr Pro Asn Gly Tyr Pro Phe Leu Phe Trp Glu Gln Tyr Ile
          995   1000   1005

Ser Leu Arg His Trp Leu Leu Leu Ser Ile Ser Val Val Leu Ala Cys
         1010   1015   1020

Thr Phe Leu Val Cys Ala Val Phe Leu Leu Asn Pro Trp Thr Ala Gly
         1025   1030   1035   1040

Ile Ile Val Met Val Leu Ala Leu Met Thr Val Glu Leu Phe Gly Met
        1045   1050   1055

Met Gly Leu Ile Gly Ile Lys Leu Ser Ala Val Pro Val Val Ile Leu
        1060   1065   1070

Ile Ala Ser Val Gly Ile Gly Val Glu Phe Thr Val His Val Ala Leu
        1075   1080   1085

Ala Phe Leu Thr Ala Ile Gly Asp Lys Asn His Arg Ala Met Leu Ala
        1090   1095   1100

Leu Glu His Met Phe Ala Pro Val Leu Asp Gly Ala Val Ser Thr Leu
1105   1110   1115   1120

Leu Gly Val Leu Met Leu Ala Gly Ser Glu Phe Asp Phe Ile Val Arg
        1125   1130   1135

Tyr Phe Phe Ala Val Leu Ala Ile Leu Thr Val Leu Gly Val Leu Asn
        1140   1145   1150

Gly Leu Val Leu Leu Pro Val Leu Leu Ser Phe Phe Gly Pro Cys Pro
        1155   1160   1165

Glu Val Ser Pro Ala Asn Gly Leu Asn Arg Leu Pro Thr Pro Ser Pro
        1170   1175   1180

Glu Pro Pro Pro Ser Val Val Arg Phe Ala Val Pro Pro Gly His Thr
1185   1190   1195   1200

Asn Asn Gly Ser Asp Ser Ser Asp Ser Glu Tyr Ser Ser Gln Thr Thr
        1205   1210   1215

Val Ser Gly Ile Ser Glu Glu Leu Arg Gln Tyr Glu Ala Gln Gln Gly
        1220   1225   1230

Ala Gly Gly Pro Ala His Gln Val Ile Val Glu Ala Thr Glu Asn Pro
        1235   1240   1245

Val Phe Ala Arg Ser Thr Val Val His Pro Asp Ser Arg His Gln Pro
        1250   1255   1260

Pro Leu Thr Pro Arg Gln Gln Pro His Leu Asp Ser Gly Ser Leu Ser
1265   1270   1275   1280

Pro Gly Arg Gln Gly Gln Gln Pro Arg Arg Asp Pro Pro Arg Glu Gly
        1285   1290   1295

Leu Arg Pro Pro Pro Tyr Arg Pro Arg Arg Asp Ala Phe Glu Ile Ser
        1300   1305   1310

Thr Glu Gly His Ser Gly Pro Ser Asn Arg Asp Arg Ser Gly Pro Arg
        1315   1320   1325

Gly Ala Arg Ser His Asn Pro Arg Asn Pro Thr Ser Thr Ala Met Gly
        1330   1335   1340

Ser Ser Val Pro Ser Tyr Cys Gln Pro Ile Thr Thr Val Thr Ala Ser
1345   1350   1355   1360

Ala Ser Val Thr Val Ala Val His Pro Pro Pro Gly Pro Gly Arg Asn
        1365   1370   1375

```
Pro Arg Gly Gly Pro Cys Pro Gly Tyr Glu Ser Tyr Pro Glu Thr Asp
        1380                    1385              1390

His Gly Val Phe Glu Asp Pro His Val Pro Phe His Val Arg Cys Glu
        1395                1400              1405

Arg Arg Asp Ser Lys Val Glu Val Ile Glu Leu Gln Asp Val Glu Cys
        1410                1415              1420

Glu Glu Arg Pro Trp Gly Ser Ser Ser Asn
1425                1430
```

(2) INFORMATION FOR SEQ ID NO:11:

        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 11 amino acids
                (B) TYPE: amino acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
Ile Ile Thr Pro Leu Asp Cys Phe Trp Glu Gly
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:12:

        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 5 amino acids
                (B) TYPE: amino acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
Leu Ile Val Gly Gly
1               5
```

(2) INFORMATION FOR SEQ ID NO:13:

        (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 7 amino acids
                (B) TYPE: amino acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: peptide

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Pro Phe Phe Trp Glu Gln Tyr
```

1   5

(2) INFORMATION FOR SEQ ID NO:14:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 28 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "primer"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGACGAATTC AARGTNCAYC ARYTNTGG   28

(2) INFORMATION FOR SEQ ID NO:15:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 26 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "primer"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGACGAATTC CYTCCCARAA RCANTC   26

(2) INFORMATION FOR SEQ ID NO:16:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 27 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid
        (A) DESCRIPTION: /desc = "primer"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:
GGACGAATTC YTNGANTGYT TYTGGGA   27

(2) INFORMATION FOR SEQ ID NO:17:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 31 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
    (A) DESCRIPTION: /desc = "primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CATACCAGCC AAGCTTGTCN GGCCARTGCA T   31

(2) INFORMATION FOR SEQ ID NO:18:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5288 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
GAATTCCGGG GACCGCAAGG AGTGCCGCGG AAGCGCCCGA AGGACAGGCT CGCTCGGCGC        60
GCCGGCTCTC GCTCTTCCGC GAACTGGATG TGGGCAGCGG CGGCCGCAGA GACCTCGGGA       120
CCCCCGCGCA ATGTGGCAAT GGAAGGCGCA GGGTCTGACT CCCCGGCAGC GGCCGCGGCC       180
GCAGCGGCAG CAGCGCCCGC CGTGTGAGCA GCAGCAGCGG CTGGTCTGTC AACCGGAGCC       240
CGAGCCCGAG CAGCCTGCGG CCAGCAGCGT CCTCGCAAGC CGAGCGCCCA GGCGCGCCAG       300
GAGCCCGCAG CAGCGGCAGC AGCGCGCCGG GCCGCCCGGG AAGCCTCCGT CCCCGCGGCG       360
GCGGCGGCGG CGGCGGCGGC AACATGGCCT CGGCTGGTAA CGCCGCCGAG CCCCAGGACC       420
GCGGCGGCGG CGGCAGCGGC TGTATCGGTG CCCCGGGACG GCCGGCTGGA GGCGGGAGGC       480
GCAGACGGAC GGGGGGGCTG CGCCGTGCTG CCGCGCCGGA CCGGGACTAT CTGCACCGGC       540
CCAGCTACTG CGACGCCGCC TTCGCTCTGG AGCAGATTTC CAAGGGGAAG GCTACTGGCC       600
GGAAAGCGCC ACTGTGGCTG AGAGCGAAGT TTCAGAGACT CTTATTTAAA CTGGGTTGTT       660
ACATTCAAAA AAACTGCGGC AAGTTCTTGG TTGTGGGCCT CCTCATATTT GGGGCCTTCG       720
CGGTGGGATT AAAAGCAGCG AACCTCGAGA CCAACGTGGA GGAGCTGTGG GTGGAAGTTG       780
GAGGACGAGT AAGTCGTGAA TTAAATTATA CTCGCCAGAA GATTGGAGAA GAGGCTATGT       840
TTAATCCTCA ACTCATGATA CAGACCCCTA AGAAGAAGG TGCTAATGTC CTGACCACAG        900
AAGCGCTCCT ACAACACCTG GACTCGGCAC TCCAGGCCAG CCGTGTCCAT GTATACATGT       960
ACAACAGGCA GTGGAAATTG GAACATTTGT GTTACAAATC AGGAGAGCTT ATCACAGAAA      1020
CAGGTTACAT GGATCAGATA ATAGAATATC TTTACCCTTG TTTGATTATT ACACCTTTGG      1080
ACTGCTTCTG GGAAGGGGCG AAATTACAGT CTGGGACAGC ATACCTCCTA GGTAAACCTC      1140
```

```
CTTTGCGGTG GACAAACTTC GACCCTTTGG AATTCCTGGA AGAGTTAAAG AAAATAAACT     1200

ATCAAGTGGA CAGCTGGGAG GAAATGCTGA ATAAGGCTGA GGTTGGTCAT GGTTACATGG     1260

ACCGCCCCTG CCTCAATCCG GCCGATCCAG ACTGCCCCGC CACAGCCCCC AACAAAAATT     1320

CAACCAAACC TCTTGATATG GCCCTTGTTT TGAATGGTGG ATGTCATGGC TTATCCAGAA     1380

AGTATATGCA CTGGCAGGAG GAGTTGATTG TGGGTGGCAC AGTCAAGAAC AGCACTGGAA     1440

AACTCGTCAG CGCCCATGCC CTGCAGACCA TGTTCCAGTT AATGACTCCC AAGCAAATGT     1500

ACGAGCACTT CAAGGGGTAC GAGTATGTCT CACACATCAA CTGGAACGAG GACAAAGCGG     1560

CAGCCATCCT GGAGGCCTGG CAGAGGACAT ATGTGGAGGT GGTTCATCAG AGTGTCGCAC     1620

AGAACTCCAC TCAAAAGGTG CTTTCCTTCA CCACCACGAC CCTGGACGAC ATCCTGAAAT     1680

CCTTCTCTGA CGTCAGTGTC ATCCGCGTGG CCAGCGGCTA CTTACTCATG CTCGCCTATG     1740

CCTGTCTAAC CATGCTGCGC TGGGACTGCT CCAAGTCCCA GGGTGCCGTG GGGCTGGCTG     1800

GCGTCCTGCT GGTTGCACTG TCAGTGGCTG CAGGACTGGG CCTGTGCTCA TTGATCGGAA     1860

TTTCCTTTAA CGCTGCAACA ACTCAGGTTT TGCCATTTCT CGCTCTTGGT GTTGGTGTGG     1920

ATGATGTTTT TCTTCTGGCC CACGCCTTCA GTGAAACAGG ACAGAATAAA AGAATCCCTT     1980

TTGAGGACAG GACCGGGGAG TGCCTGAAGC GCACAGGAGC CAGCGTGGCC CTCACGTCCA     2040

TCAGCAATGT CACAGCCTTC TTCATGGCCG CGTTAATCCC AATTCCCGCT CTGCGGGCGT     2100

TCTCCCTCCA GGCAGCGGTA GTAGTGGTGT TCAATTTTGC CATGGTTCTG CTCATTTTTC     2160

CTGCAATTCT CAGCATGGAT TTATATCGAC GCGAGGACAG GAGACTGGAT ATTTTCTGCT     2220

GTTTTACAAG CCCCTGCGTC AGCAGAGTGA TTCAGGTTGA ACCTCAGGCC TACACCGACA     2280

CACACGACAA TACCCGCTAC AGCCCCCCAC CTCCCTACAG CAGCCACAGC TTTGCCCATG     2340

AAACGCAGAT TACCATGCAG TCCACTGTCC AGCTCCGCAC GGAGTACGAC CCCCACACGC     2400

ACGTGTACTA CACCACCGCT GAGCCGCGCT CCGAGATCTC TGTGCAGCCC GTCACCGTGA     2460

CACAGGACAC CCTCAGCTGC CAGAGCCCAG AGAGCACCAG CTCCACAAGG GACCTGCTCT     2520

CCCAGTTCTC CGACTCCAGC CTCCACTGCC TCGAGCCCCC CTGTACGAAG TGGACACTCT     2580

CATCTTTTGC TGAGAAGCAC TATGCTCCTT TCCTCTTGAA ACCAAAAGCC AAGGTAGTGG     2640

TGATCTTCCT TTTTCTGGGC TTGCTGGGGG TCAGCCTTTA TGGCACCACC CGAGTGAGAG     2700

ACGGGCTGGA CCTTACGGAC ATTGTACCTC GGGAAACCAG AGAATATGAC TTTATTGCTG     2760

CACAATTCAA ATACTTTTCT TTCTACAACA TGTATATAGT CACCCAGAAA GCAGACTACC     2820

CGAATATCCA GCACTTACTT TACGACCTAC ACAGGAGTTT CAGTAACGTG AAGTATGTCA     2880

TGTTGGAAGA AAACAAACAG CTTCCCAAAA TGTGGCTGCA CTACTTCAGA GACTGGCTTC     2940

ACTCACTTCA GGATGCATTT GACAGTGACT GGGAAACCGG GAAAATCATG CCAAACAATT     3000

ACAAGAATGG ATCAGACGAT GGAGTCCTTG CCTACAAACT CCTGGTGCAA ACCGGCAGCC     3060
```

72

```
GCGATAAGCC CATCGACATC AGCCAGTTGA CTAAACAGCG TCTGGTGGAT GCAGATGGCA    3120
TCATTAATCC CAGCGCTTTC TACATCTACC TGACGGCTTG GGTCAGCAAC GACCCCGTCG    3180
CGTATGCTGC CTCCCAGGCC AACATCCGGC CACACCGACC AGAATGGGTC CACGACAAAG    3240
CCGACTACAT GCCTGAAACA AGGCTGAGAA TCCCGGCAGC AGAGCCCATC GAGTATGCCC    3300
AGTTCCCTTT CTACCTCAAC GGGTTGCGGG ACACCTCAGA CTTTGTGGAG GCAATTGAAA    3360
AAGTAAGGAC CATCTGCAGC AACTATACGA GCCTGGGGCT GTCCAGTTAC CCCAACGGCT    3420
ACCCCTTCCT CTTCTGGGAG CAGTACATCG GCCTCCGCCA CTGGCTGCTG CTGTTCATCA    3480
GCGTGGTGTT GGCCTGCACA TTCCTCGTGT GCGCTGTCTT CCTTCTGAAC CCCTGGACGG    3540
CCGGGATCAT TGTGATGGTC CTGGCGCTGA TGACGGTCGA GCTGTTCGGC ATGATGGGCC    3600
TCATCGGAAT CAAGCTCAGT GCCGTGCCCG TGGTCATCCT GATCGCTTCT GTTGGCATAG    3660
GAGTGGAGTT CACCGTTCAC GTTGCTTTGG CCTTTCTGAC GGCCATCGGC GACAAGAACC    3720
GCAGGGCTGT GCTTGCCCTG GAGCACATGT TTGCACCCGT CCTGGATGGC GCCGTGTCCA    3780
CTCTGCTGGG AGTGCTGATG CTGGCGGGAT CTGAGTTCGA CTTCATTGTC AGGTATTTCT    3840
TTGCTGTGCT GGCGATCCTC ACCATCCTCG GCGTTCTCAA TGGGCTGGTT TTGCTTCCCG    3900
TGCTTTTGTC TTTCTTTGGA CCATATCCTG AGGTGTCTCC AGCCAACGGC TTGAACCGCC    3960
TGCCCACACC CTCCCCTGAG CCACCCCCCA GCGTGGTCCG CTTCGCCATG CCGCCCGGCC    4020
ACACGCACAG CGGGTCTGAT TCCTCCGACT CGGAGTATAG TTCCCAGACG ACAGTGTCAG    4080
GCCTCAGCGA GGAGCTTCGG CACTACGAGG CCCAGCAGGG CGCGGGAGGC CCTGCCCACC    4140
AAGTGATCGT GGAAGCCACA GAAAACCCCG TCTTCGCCCA CTCCACTGTG GTCCATCCCG    4200
AATCCAGGCA TCACCCACCC TCGAACCCGA GACAGCAGCC CCACCTGGAC TCAGGGTCCC    4260
TGCCTCCCGG ACGGCAAGGC CAGCAGCCCC GCAGGGACCC CCCCAGAGAA GGCTTGTGGC    4320
CACCCCTCTA CAGACCGCGC AGAGACGCTT TTGAAATTTC TACTGAAGGG CATTCTGGCC    4380
CTAGCAATAG GGCCCGCTGG GGCCCTCGCG GGGCCCGTTC TCACAACCCT CGGAACCCAG    4440
CGTCCACTGC CATGGGCAGC TCCGTGCCCG GCTACTGCCA GCCCATCACC ACTGTGACGG    4500
CTTCTGCCTC CGTGACTGTC GCCGTGCACC CGCCGCCTGT CCCTGGGCCT GGGCGGAACC    4560
CCCGAGGGGG ACTCTGCCCA GGCTACCCTG AGACTGACCA CGGCCTGTTT GAGGACCCCC    4620
ACGTGCCTTT CCACGTCCGG TGTGAGAGGA GGGATTCGAA GGTGGAAGTC ATTGAGCTGC    4680
AGGACGTGGA ATGCGAGGAG AGGCCCCGGG GAAGCAGCTC CAACTGAGGG TGATTAAAAT    4740
CTGAAGCAAA GAGGCCAAAG ATTGGAAACC CCCCACCCCC ACCTCTTTCC AGAACTGCTT    4800
GAAGAGAACT GGTTGGAGTT ATGGAAAAGA TGCCCTGTGC CAGGACAGCA GTTCATTGTT    4860
ACTGTAACCG ATTGTATTAT TTTGTTAAAT ATTTCTATAA ATATTTAAGA GATGTACACA    4920
```

```
TGTGTAATAT AGGAAGGAAG GATGTAAAGT GGTATGATCT GGGGCTTCTC CACTCCTGCC      4980

CCAGAGTGTG GAGGCCACAG TGGGGCCTCT CCGTATTTGT GCATTGGGCT CCGTGCCACA      5040

ACCAAGCTTC ATTAGTCTTA AATTTCAGCA TATGTTGCTG CTGCTTAAAT ATTGTATAAT      5100

TTACTTGTAT AATTCTATGC AAATATTGCT TATGTAATAG GATTATTTTG TAAAGGTTTC      5160

TGTTTAAAAT ATTTTAAATT TGCATATCAC AACCCTGTGG TAGTATGAAA TGTTACTGTT      5220

AACTTTCAAA CACGCTATGC GTGATAATTT TTTTGTTTAA TGAGCAGATA TGAAGAAAGC      5280

CCGGAATT                                                                5288
```

(2) INFORMATION FOR SEQ ID NO:19:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1447 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
Met Ala Ser Ala Gly Asn Ala Ala Glu Pro Gln Asp Arg Gly Gly Gly
1               5                   10                  15

Gly Ser Gly Cys Ile Gly Ala Pro Gly Arg Pro Ala Gly Gly Gly Arg
            20              25                  30

Arg Arg Arg Thr Gly Gly Leu Arg Arg Ala Ala Ala Pro Asp Arg Asp
        35                  40                  45

Tyr Leu His Arg Pro Ser Tyr Cys Asp Ala Ala Phe Ala Leu Glu Gln
    50              55                  60

Ile Ser Lys Gly Lys Ala Thr Gly Arg Lys Ala Pro Leu Trp Leu Arg
65              70                  75                  80

Ala Lys Phe Gln Arg Leu Leu Phe Lys Leu Gly Cys Tyr Ile Gln Lys
            85                  90                  95

Asn Cys Gly Lys Phe Leu Val Val Gly Leu Leu Ile Phe Gly Ala Phe
            100                 105                 110

Ala Val Gly Leu Lys Ala Ala Asn Leu Glu Thr Asn Val Glu Glu Leu
            115                 120                 125

Trp Val Glu Val Gly Gly Arg Val Ser Arg Glu Leu Asn Tyr Thr Arg
    130                 135                 140

Gln Lys Ile Gly Glu Glu Ala Met Phe Asn Pro Gln Leu Met Ile Gln
145                 150                 155                 160

Thr Pro Lys Glu Glu Gly Ala Asn Val Leu Thr Thr Glu Ala Leu Leu
                165                 170                 175

Gln His Leu Asp Ser Ala Leu Gln Ala Ser Arg Val His Val Tyr Met
```

```
                180                    185                    190
Tyr Asn Arg Gln Trp Lys Leu Glu His Leu Cys Tyr Lys Ser Gly Glu
        195                 200                 205

Leu Ile Thr Glu Thr Gly Tyr Met Asp Gln Ile Ile Glu Tyr Leu Tyr
        210                 215                 220

Pro Cys Leu Ile Ile Thr Pro Leu Asp Cys Phe Trp Glu Gly Ala Lys
225                     230                 235                 240

Leu Gln Ser Gly Thr Ala Tyr Leu Leu Gly Lys Pro Pro Leu Arg Trp
                245                 250                 255

Thr Asn Phe Asp Pro Leu Glu Phe Leu Glu Glu Leu Lys Lys Ile Asn
                260                 265                 270

Tyr Gln Val Asp Ser Trp Glu Glu Met Leu Asn Lys Ala Glu Val Gly
                275                 280                 285

His Gly Tyr Met Asp Arg Pro Cys Leu Asn Pro Ala Asp Pro Asp Cys
        290                 295                 300

Pro Ala Thr Ala Pro Asn Lys Asn Ser Thr Lys Pro Leu Asp Met Ala
305                     310                 315                 320

Leu Val Leu Asn Gly Gly Cys His Gly Leu Ser Arg Lys Tyr Met His
                325                 330                 335

Trp Gln Glu Glu Leu Ile Val Gly Gly Thr Val Lys Asn Ser Thr Gly
                340                 345                 350

Lys Leu Val Ser Ala His Ala Leu Gln Thr Met Phe Gln Leu Met Thr
                355                 360                 365

Pro Lys Gln Met Tyr Glu His Phe Lys Gly Tyr Glu Tyr Val Ser His
        370                 375                 380

Ile Asn Trp Asn Glu Asp Lys Ala Ala Ala Ile Leu Glu Ala Trp Gln
385                     390                 395                 400

Arg Thr Tyr Val Glu Val Val His Gln Ser Val Ala Gln Asn Ser Thr
                405                 410                 415

Gln Lys Val Leu Ser Phe Thr Thr Thr Thr Leu Asp Asp Ile Leu Lys
                420                 425                 430

Ser Phe Ser Asp Val Ser Val Ile Arg Val Ala Ser Gly Tyr Leu Leu
        435                 440                 445

Met Leu Ala Tyr Ala Cys Leu Thr Met Leu Arg Trp Asp Cys Ser Lys
        450                 455                 460

Ser Gln Gly Ala Val Gly Leu Ala Gly Val Leu Leu Val Ala Leu Ser
465                 470                 475                 480

Val Ala Ala Gly Leu Gly Leu Cys Ser Leu Ile Gly Ile Ser Phe Asn
                485                 490                 495

Ala Ala Thr Thr Gln Val Leu Pro Phe Leu Ala Leu Gly Val Gly Val
        500                 505                 510

Asp Asp Val Phe Leu Leu Ala His Ala Phe Ser Glu Thr Gly Gln Asn
```

```
                   515                      520                      525

   Lys Arg Ile Pro Phe Glu Asp Arg Thr Gly Glu Cys Leu Lys Arg Thr
       530                 535                 540

   Gly Ala Ser Val Ala Leu Thr Ser Ile Ser Asn Val Thr Ala Phe Phe
   545                 550                 555                 560

   Met Ala Ala Leu Ile Pro Ile Pro Ala Leu Arg Ala Phe Ser Leu Gln
                   565                 570                 575

   Ala Ala Val Val Val Val Phe Asn Phe Ala Met Val Leu Leu Ile Phe
                   580                 585                 590

   Pro Ala Ile Leu Ser Met Asp Leu Tyr Arg Arg Glu Asp Arg Arg Leu
               595                 600                 605

   Asp Ile Phe Cys Cys Phe Thr Ser Pro Cys Val Ser Arg Val Ile Gln
       610                 615                 620

   Val Glu Pro Gln Ala Tyr Thr Asp Thr His Asp Asn Thr Arg Tyr Ser
   625                 630                 635                 640

   Pro Pro Pro Pro Tyr Ser Ser His Ser Phe Ala His Glu Thr Gln Ile
                   645                 650                 655

   Thr Met Gln Ser Thr Val Gln Leu Arg Thr Glu Tyr Asp Pro His Thr
               660                 665                 670

   His Val Tyr Tyr Thr Thr Ala Glu Pro Arg Ser Glu Ile Ser Val Gln
       675                     680                 685

   Pro Val Thr Val Thr Gln Asp Thr Leu Ser Cys Gln Ser Pro Glu Ser
       690                 695                 700

   Thr Ser Ser Thr Arg Asp Leu Leu Ser Gln Phe Ser Asp Ser Ser Leu
   705                 710                 715                 720

   His Cys Leu Glu Pro Pro Cys Thr Lys Trp Thr Leu Ser Ser Phe Ala
                   725                 730                 735

   Glu Lys His Tyr Ala Pro Phe Leu Leu Lys Pro Lys Ala Lys Val Val
               740                 745                 750

   Val Ile Phe Leu Phe Leu Gly Leu Leu Gly Val Ser Leu Tyr Gly Thr
               755                 760                 765

   Thr Arg Val Arg Asp Gly Leu Asp Leu Thr Asp Ile Val Pro Arg Glu
       770                 775                 780

   Thr Arg Glu Tyr Asp Phe Ile Ala Ala Gln Phe Lys Tyr Phe Ser Phe
   785                 790                 795                 800

   Tyr Asn Met Tyr Ile Val Thr Gln Lys Ala Asp Tyr Pro Asn Ile Gln
                   805                 810                 815

   His Leu Leu Tyr Asp Leu His Arg Ser Phe Ser Asn Val Lys Tyr Val
               820                 825                 830

   Met Leu Glu Glu Asn Lys Gln Leu Pro Lys Met Trp Leu His Tyr Phe
               835                 840                 845

   Arg Asp Trp Leu Gln Gly Leu Gln Asp Ala Phe Asp Ser Asp Trp Glu
       850                 855                 860
```

Thr Gly Lys Ile Met Pro Asn Asn Tyr Lys Asn Gly Ser Asp Asp Gly
865   870   875   880

Val Leu Ala Tyr Lys Leu Leu Val Gln Thr Gly Ser Arg Asp Lys Pro
               885   890   895

Ile Asp Ile Ser Gln Leu Thr Lys Gln Arg Leu Val Asp Ala Asp Gly
               900   905   910

Ile Ile Asn Pro Ser Ala Phe Tyr Ile Tyr Leu Thr Ala Trp Val Ser
          915   920   925

Asn Asp Pro Val Ala Tyr Ala Ala Ser Gln Ala Asn Ile Arg Pro His
     930   935   940

Arg Pro Glu Trp Val His Asp Lys Ala Asp Tyr Met Pro Glu Thr Arg
945   950   955   960

Leu Arg Ile Pro Ala Ala Glu Pro Ile Glu Tyr Ala Gln Phe Pro Phe
               965   970   975

Tyr Leu Asn Gly Leu Arg Asp Thr Ser Asp Phe Val Glu Ala Ile Glu
          980   985   990

Lys Val Arg Thr Ile Cys Ser Asn Tyr Thr Ser Leu Gly Leu Ser Ser
     995   1000   1005

Tyr Pro Asn Gly Tyr Pro Phe Leu Phe Trp Glu Gln Tyr Ile Gly Leu
     1010   1015   1020

Arg His Trp Leu Leu Leu Phe Ile Ser Val Val Leu Ala Cys Thr Phe
1025   1030   1035   1040

Leu Val Cys Ala Val Phe Leu Leu Asn Pro Trp Thr Ala Gly Ile Ile
               1045   1050   1055

Val Met Val Leu Ala Leu Met Thr Val Glu Leu Phe Gly Met Met Gly
          1060   1065   1070

Leu Ile Gly Ile Lys Leu Ser Ala Val Pro Val Val Ile Leu Ile Ala
          1075   1080   1085

Ser Val Gly Ile Gly Val Glu Phe Thr Val His Val Ala Leu Ala Phe
     1090   1095   1100

Leu Thr Ala Ile Gly Asp Lys Asn Arg Arg Ala Val Leu Ala Leu Glu
1105   1110   1115   1120

His Met Phe Ala Pro Val Leu Asp Gly Ala Val Ser Thr Leu Leu Gly
               1125   1130   1135

Val Leu Met Leu Ala Gly Ser Glu Phe Asp Phe Ile Val Arg Tyr Phe
          1140   1145   1150

Phe Ala Val Leu Ala Ile Leu Thr Ile Leu Gly Val Leu Asn Gly Leu
          1155   1160   1165

Val Leu Leu Pro Val Leu Leu Ser Phe Phe Gly Pro Tyr Pro Glu Val
     1170   1175   1180

Ser Pro Ala Asn Gly Leu Asn Arg Leu Pro Thr Pro Ser Pro Glu Pro
1185   1190   1195   1200

```
Pro Pro Ser Val Val Arg Phe Ala Met Pro Pro Gly His Thr His Ser
            1205                1210            1215

Gly Ser Asp Ser Ser Asp Ser Glu Tyr Ser Ser Gln Thr Thr Val Ser
            1220                1225            1230

Gly Leu Ser Glu Glu Leu Arg His Tyr Glu Ala Gln Gln Gly Ala Gly
            1235            1240                1245

Gly Pro Ala His Gln Val Ile Val Glu Ala Thr Glu Asn Pro Val Phe
        1250            1255            1260

Ala His Ser Thr Val Val His Pro Glu Ser Arg His His Pro Pro Ser
1265            1270            1275            1280

Asn Pro Arg Gln Gln Pro His Leu Asp Ser Gly Ser Leu Pro Pro Gly
                1285            1290            1295

Arg Gln Gly Gln Gln Pro Arg Arg Asp Pro Pro Arg Glu Gly Leu Trp
            1300            1305            1310

Pro Pro Leu Tyr Arg Pro Arg Arg Asp Ala Phe Glu Ile Ser Thr Glu
            1315            1320            1325

Gly His Ser Gly Pro Ser Asn Arg Ala Arg Trp Gly Pro Arg Gly Ala
        1330            1335            1340

Arg Ser His Asn Pro Arg Asn Pro Ala Ser Thr Ala Met Gly Ser Ser
1345            1350            1355            1360

Val Pro Gly Tyr Cys Gln Pro Ile Thr Thr Val Thr Ala Ser Ala Ser
            1365            1370            1375

Val Thr Val Ala Val His Pro Pro Pro Val Pro Gly Pro Gly Arg Asn
            1380            1385            1390

Pro Arg Gly Gly Leu Cys Pro Gly Tyr Pro Glu Thr Asp His Gly Leu
        1395            1400            1405

Phe Glu Asp Pro His Val Pro Phe His Val Arg Cys Glu Arg Arg Asp
        1410            1415            1420

Ser Lys Val Glu Val Ile Glu Leu Gln Asp Val Glu Cys Glu Glu Arg
1425            1430            1435            1440

Pro Arg Gly Ser Ser Ser Asn
            1445
```